

Trabajo de Fin de Grado

Diseño e implementación de un módulo para la
anotación de imágenes médicas

Isaac Sánchez Ruiz (s10m024),
Grado en Matemáticas e Informática,
Escuela Técnica Superior de Ingenieros Informáticos (ETSIINF),
Universidad Politécnica de Madrid

6 de junio de 2014

Índice general

1. Introducción	1
2. Objetivos	6
2.1. Visión general del proyecto	6
2.2. Objetivos específicos de este proyecto de Fin de Grado	8
3. Estado del arte	10
3.1. Conceptos previos	10
3.1.1. Segmentación de imágenes	10
3.1.2. Proceso de segmentación	12
3.1.3. Características de las imágenes PET y CT	13
3.2. Técnicas principales de segmentación	18
3.3. Segmentación manual y semiautomática	26
3.4. Segmentación automática	28
3.4.1. Técnicas de umbralización	29
3.4.2. Técnicas de multiumbralización	37
4. Metodología	40
4.1. Técnicas escogidas	40
4.1.1. Umbralización de Otsu	42
4.1.2. Umbralización de Ridler-Calvard	45
4.1.3. Umbral fijo de Paulino	46
4.1.4. Umbralización fija de Erdi	47
4.2. Descripción del material de prueba	48
4.2.1. Formato de las imágenes PET	48
4.2.2. Características de los pacientes	49
4.3. Métodos de evaluación	53
5. Resultados	56
6. Conclusión	66

Índice de figuras

2.1. Esquema de la visión general del proyecto	7
3.1. Ejemplo de imagen CT de tórax	13
3.2. Ejemplo de imagen PET de paciente con cáncer de pulmón	14
3.3. Aplicación de visualización PET-CT [5]	16
3.4. Ejemplo de híbrido PET-CT 3D	17
3.5. Ejemplo de región sin detectar en CT pero visible en PET [5]	18
3.6. Proceso descrito de <i>hill climbing</i> [1]	22
3.7. Ejemplo de algoritmo de Watershed descrito [1]	23
3.8. (A) imagen original, (B) tras aplicar el operador Sobel, (C y D) tras umbralizar con distintos umbrales [1]	24
3.9. Diferencias en 3 tumores distintos entre varios observadores indicado mediante contornos azules [12]	27
3.10. Diferentes resultados obtenidos en el mismo paciente entre distintas semillas iniciales (<i>slices</i> 79-85) [8]	28
3.11. Ejemplos de imágenes umbralizadas en NDT [20] y binarización de documentos	33
3.12. Esferas 1, 3 y 5 en ambos escáneres y los niveles de contraste [17]	34
3.13. Medida de rendimiento de las 12 técnicas probadas [17]	35
4.1. Ejemplos de histogramas bimodales	43
4.2. Segmentación inicial [18]	45
4.3. Esquema del algoritmo de Ridler [18]	46
4.4. Volúmenes 1 (rojo), 2 (amarillo) e intersección (verde)	51
4.5. Volúmenes 5 y 6	51
4.6. Reconstrucción 3D de las imágenes PET	52
5.1. Resultados en varios pacientes	64
5.2. <i>Slice</i> 85 del paciente 3	65

Índice de cuadros

4.1. Tabla de SUV máximo	49
4.2. Tabla de slices afectados de los pacientes	50
5.1. Umbrales obtenidos con cada técnica sobre el SUV	56
5.2. Resultados de evaluación usando la técnica de Otsu	58
5.3. Evaluación de casos especiales usando Otsu	58
5.4. Resultados de evaluación usando la técnica de Ridler	59
5.5. Evaluación de casos especiales usando Ridler	59
5.6. Resultados de evaluación usando la técnica de Paulino	60
5.7. Evaluación de casos especiales usando Paulino	60
5.8. Resultados de evaluación usando la técnica de Erdi	61
5.9. Evaluación de casos especiales usando Erdi	61
5.10. Datos de volumen, SUV mínimo y máximo en cada una de las ROIs manuales	63

Resumen

Desde hace tiempo ha habido mucho interés en la automatización de todo tipo de tareas en las que la intervención humana es esencial para que sean completadas con éxito. Esto es de especial interés si además se necesita de personal muy cualificado para ciertas tareas que pueden ser perfectamente reproducibles y, o bien requieren mucha formación, o bien consumen mucho tiempo.

Este proyecto está dirigido a la búsqueda de métodos para automatizar la anotación de imágenes médicas. En concreto, se centra en el apartado de delimitación de las regiones de interés (ROIs) en imágenes de tipo PET siendo éstas usadas con frecuencia junto con las imágenes de tipo CT en el campo de oncología para delinear volúmenes afectados por cáncer.

Se pretende con esto ayudar a los hospitales a organizar y estructurar las imágenes de sus pacientes y relacionarlas con las notas clínicas. Esto es lo que llamaremos el proceso de anotación de imágenes y la integración con la anotación de notas clínicas respectivamente.

En este documento nos vamos a centrar en describir cuáles eran los objetivos iniciales, los pasos dados para su consecución y las dificultades encontradas durante el proceso.

De todas las técnicas existentes en la literatura, se han elegido 4 técnicas de segmentación, 2 de ellas probadas en pacientes reales y las otras 2 probadas sólo en *phantoms* según la literatura. En nuestro caso, las pruebas, se han realizado en imágenes PET de 6 pacientes reales diagnosticados de cáncer. Los resultados han sido analizados y presentados.

Abstract

For a long period of time, there has been an increasing interest in automation of tasks where human intervention is needed in order to succeed. This interest is even greater if those tasks must be solved by qualified specialists in the area and the task is reproducible or if the task is too time consuming.

The main objective of this project is to find methods which can help to automate medical image annotation processes. In our specific case, we are willing to delineate regions of interest (ROIs) in PET images which are frequently used simultaneously with CT images in oncology to determine those volumes that are affected by cancer.

With this process we want to help hospitals organize and structure the images they have from their patient studies and to relate these images to the corpus annotations. We may call this the image annotation process and the integration with the corpus annotation respectively.

In this document we are going to concentrate in the description of the initial objectives, the steps we had to go through and the difficulties we had to face during this process.

From all existing techniques in the literature, 4 segmentation techniques have been chosen, 2 of them were tested in real patients and the other 2 were tested using *phantoms* according to the literature. In our case, the tests have been done using PET images from 6 real patients diagnosed with cancer. The results have been analyzed and presented.

Capítulo 1

Introducción

En la actualidad, el uso de imágenes médicas está muy extendido en el campo de la oncología. El uso de imágenes en el diagnóstico y tratamiento de cáncer ayuda a los especialistas a llevar a cabo sus tareas de forma más rápida. Esto es vital en el diagnóstico de pacientes que puedan padecer cáncer ya que la detección temprana de esta enfermedad puede ayudar a llevar a cabo un tratamiento rápido y en consecuencia, reducir los daños que pueda causar e incluso evitar el fallecimiento del paciente.

Las técnicas de obtención de imágenes y los distintos tipos de imágenes médicas que se usan en la actualidad es muy variado. Cada técnica proporciona un tipo de imágenes con características distintas. Por su naturaleza, ciertas características se adaptan mejor a ciertos tipos de problemas y, por tanto, el uso de las técnicas está más extendido en unas áreas dentro de oncología que en otras.

De las regiones cancerígenas detectadas en las imágenes, se suelen extraer importantes datos como la forma, el volumen, la actividad metabólica y la estructura de dicho volumen. La correcta extracción de estos datos depende en mayor o menor medida del tipo de imagen que se esté utilizando.

Por esta razón, la combinación de la información contenida en distintas imágenes es una tendencia cada vez más popular. Así, se pretende aprovechar las ventajas que proporcionan los diversos tipos de imágenes y superar, al ser posible, las restricciones que tiene el uso de una sola técnica.

Los tipos de imágenes más comunes en oncología se obtienen a través de la interacción de ondas electromagnéticas con los tejidos y fluidos de los pacientes. Aunque

también existe otra técnica para obtener imágenes que hace uso de ondas acústicas. Esta técnica se conoce como ultrasonido y suele tener aplicación en el diagnóstico de cáncer de mama, próstata y fibrosis del hígado [6]. También son comunes las imágenes de resonancia magnética (MRI) las cuales se usan para detectar cáncer, para estudiar la respuesta a terapia y guiar la terapia de forma mínimamente invasiva, entre otros.

Las características de las imágenes obtenidas a partir de la radiación electromagnética varían dependiendo de la frecuencia de la onda electromagnética. En el espectro de baja frecuencia (entre radiofrecuencia y luz visible), se encuentran técnicas de obtención de imágenes como tomografías cerca del infrarrojo (NIR), imágenes termo- y foto-acústicas; y tomografía de impedancia eléctrica (EIT). Si bien este tipo de imágenes es de utilidad, no son tan comunes como las imágenes de mayor frecuencia ni como las de ultrasonidos, que han sido mencionadas en el párrafo anterior.

Entre las imágenes de alta frecuencia y las de baja frecuencia, se encuentran las imágenes generadas mediante luz visible. Éstas se conocen como imágenes de tipo óptico y han quedado limitadas al uso en endoscopias, catéteres e imágenes superficiales [6].

Por último, las imágenes de alta frecuencia (por encima de la frecuencia del ultravioleta) son muy comunes y son las que se consideran de tipo ionizante. Imágenes de este tipo son las imágenes basadas en rayos X, las imágenes de medicina nuclear y aquellas basadas en emisión de positrones (PET). De todos estos tipos de imágenes existen diversos derivados que tienen utilidad en diferentes áreas. Este tipo de imágenes son las que se usan en este trabajo, concretamente las imágenes de tipo PET.

Todas las imágenes que hacen uso de radiación de baja frecuencia y aquellas que hacen uso del espectro de luz visible, se dice que usan radiación no ionizante, es decir, que la radiación emitida no es capaz de liberar electrones de átomos y moléculas, por lo que se dice que no los ioniza. La radiación ionizante supone un peligro para el paciente puesto que puede destruir y llegar a generar cáncer en tejidos sanos tras un largo período de exposición a una fuente de radiación ionizante.

Por otro lado, se están haciendo avances usando distintos tipos de marcadores que parecen prometedores para identificar tumores cancerígenos. Últimamente, ha aumentado el uso de marcadores en las imágenes médicas para resaltar y mostrar los efectos del metabolismo, la proliferación y migración de células entre otros [6].

Con esto, hemos introducido los tipos básicos de imágenes y su capacidad, centrándonos ahora en las imágenes de tipo CT, de PET y, posteriormente, el híbrido PET-CT.

El escáner de imágenes CT (Computed Tomography) está clasificado dentro de las

imágenes que utilizan rayos X. Sin embargo, son distintos de las pruebas de rayos X a las que estamos acostumbrados en casos comunes de traumatismo tales como fracturas o fisuras de huesos. La diferencia reside en que las imágenes CT no son planas, suelen ser tridimensionales o incluso cuatro-dimensionales. Los casos cuatro dimensionales son aquellos en los que se incluye la dimensión del tiempo. Se utiliza en casos en los que haya que observar el comportamiento del tumor tal y como sucede en pacientes con cáncer de pulmón. En estos casos, permite a los especialistas observar el tumor durante todo el ciclo de respiración del paciente. Este tipo de imágenes se conocen como imágenes *anatómicas* debido a su naturaleza tal y como sucede con las imágenes de rayos X.

Las imágenes de tipo CT se pueden tomar principalmente con dos tipos de escáneres: el c-arm CT que está basado en un brazo que gira alrededor del paciente obteniendo varias imágenes desde distintas perspectivas para la reconstrucción 3D [14]; y el multi-slice CT que se basa en grandes áreas de detectores usando tubos de rayos X de alta potencia, de esta forma se pueden cubrir grandes volúmenes de tórax y abdomen mientras el paciente aguanta la respiración. También se ha utilizado éste último para escanear el cerebro [6].

Las técnicas de CAD (Computer Aided Diagnostics) usando imágenes CT se han extendido sobre todo para cáncer de pulmón y de colon. En los casos de cáncer de pulmón se identifican a menudo nódulos sanos de los pulmones como cancerígenos. Esto sucede porque el 50 % de las ocasiones, los nódulos son tumores benignos y son casi indistinguibles en CT [6].

En resumen, las imágenes CT tienen el beneficio de las imágenes de rayos X, es decir, una muy alta resolución debido a su naturaleza, pero tienen muy poca sensibilidad y como hemos visto antes, son de tipo ionizante. Y en este tipo de imágenes, los tejidos blandos tienden a ser difíciles de distinguir los unos de los otros tal y como sucede en los nódulos pulmonares.

En cuanto a las imágenes PET (Positron-Emission Tomography), es un tipo de tomografía que se obtiene tras haber emitido positrones mediante radioisotopos como ^{18}F , ^{11}C y ^{68}Ga entre otros. El que se utiliza normalmente es el ^{18}F . De hecho, sólo hay dos marcadores aprobados y ambos hacen uso del ^{18}F . Éstos son el ^{18}F -FDG que contiene glucosa y se puede usar para marcar el metabolismo; y el ^{18}F -NaF₂ que se utiliza para ser incorporado en los huesos [6]. Este tipo de imágenes se clasifican como imágenes *metabólicas*.

Las imágenes PET se suelen utilizar para distinguir volúmenes de tejido afectado por cáncer de aquéllos que no lo están. Al principio, se decidió utilizar el volumen del tumor como medida de desarrollo del cáncer. Pero usar solamente el volumen de los

tumores como medida de la evolución de la enfermedad era inadecuado ya que sólo puede mostrar un retraso en la respuesta a la terapia y no da indicaciones sobre el metabolismo y otros parámetros. Ésta es la principal razón del uso de los marcadores en este tipo de imágenes [6].

En este trabajo, nos vamos a centrar en imágenes de FDG PET, que usan el primer marcador. Este marcador destacará los tumores malignos en la imagen ya que éstos tienden a tener un metabolismo mucho mayor que el de los demás tipos de tejidos. Por tanto, cuando las células cancerígenas consumen el FDG de la sustancia y hay una mayor concentración del radioisótopo ^{18}F en dicha zona. Éste emite positrones que son detectados por el escáner y así se acaba generando la imagen con mayor brillo o intensidad en torno a esa zona.

Las principales aplicaciones de las imágenes PET incluyen diagnóstico del tumor, estadificación¹, detección de recaída, seguimiento y evaluación de respuesta ante la terapia [17]. Las zonas del cuerpo donde las imágenes PET no son útiles son aquellas áreas donde existe una concentración del marcador que no se corresponde con un tumor maligno. Estas áreas comprenden el cerebro, que siempre tiene un metabolismo alto; los riñones, la próstata y la vejiga [6].

En radioterapia, es muy importante conocer la extensión de los tumores malignos y para ello es necesario conocer el volumen de estos tumores. El volumen se puede determinar usando imágenes de tipo PET. El procedimiento incluye la delineación de la zona afectada por parte de un experto y del resultado obtenido, se obtiene el volumen estimado del tumor. Los volúmenes afectados se conocen como regiones de interés (ROIs) aunque este término se usa también en otras áreas de medicina [17].

Debido a la naturaleza de las imágenes PET, la tarea de delineación manual no es tan sencilla como pueda ser esa misma tarea en otro tipo de volúmenes en CT tal y como sucede con la distinción de huesos y tejidos blandos. Esto se debe a que la resolución espacial de las imágenes PET y el nivel de ruido es relativamente mayor que el existente en las imágenes CT. Por tanto, la determinación precisa de la zona cancerígena y la no cancerígena es una tarea complicada.

Por este motivo, la combinación de imágenes PET y CT es una buena alternativa en el sentido de que se pueden combinar las ventajas de ambos tipos de imágenes. Estas ventajas son la sensibilidad de las imágenes PET usando el marcador ^{18}F -FDG para la correcta evaluación del progreso del tumor cancerígeno y la resolución espacial que proporcionan las imágenes CT. Así las tareas de diagnóstico y pronóstico se pueden

¹Se conoce como el proceso de determinación del nivel de extensión del cáncer en el paciente.

realizar con mayor precisión.

Este tipo de imagen se conoce como PET-CT y se consigue realizando las pruebas de PET y CT al mismo tiempo sobre el paciente de forma que no haya variaciones entre ambos tipos de imágenes. Si no se hiciesen simultáneamente, se correría un alto riesgo de que las dos imágenes no quedasen alineadas y carecería de sentido combinarlas.

A lo largo de este documento, se revisarán las técnicas encontradas en diversos artículos para delinear las regiones de interés de las imágenes PET, es decir, los volúmenes afectados por cáncer. Esta revisión se hará en el capítulo 3 del estado del arte.

En el capítulo 4 de metodología, se procederá a describir las técnicas escogidas para realizar las pruebas de delineación en los pacientes de los que se dispone para este proyecto así como la descripción de las características del material de prueba.

También se hará un análisis de los resultados obtenidos en el capítulo 5 con los métodos elegidos y con el material de prueba ya descrito y, por último, en el capítulo 6 se dará una conclusión relacionando los resultados y se hará una reflexión para que futuros interesados puedan continuar con el desarrollo de este tema.

Capítulo 2

Objetivos

2.1. Visión general del proyecto

Este proyecto Fin de Grado se enmarca en uno de los proyectos de investigación que se están desarrollando en el Laboratorio de Análisis de Datos y Simulación (MIDAS) del Centro de Tecnología Biomédica de la UPM. El objetivo de este proyecto es la anotación de imágenes médicas junto con la anotación de notas clínicas.

El objetivo general de dicho proyecto es conseguir una herramienta que pueda ayudar a los médicos de los hospitales actuales y, por tanto, con el material actual; a estructurar, la información electrónica disponible de los pacientes de manera automática, para su posterior almacenamiento y consulta. Esta herramienta o herramientas están idealmente compuestas de varios módulos que deberían ser capaces de coordinarse entre sí para conseguir el objetivo anterior.

El proyecto consta entonces de las siguientes partes desde un punto de vista general:

- Anotación de imágenes
 - Delineación de las regiones de interés (ROI)
 - Delineación a baja resolución espacial usando imágenes PET
 - Delineación a alta resolución espacial usando imágenes CT a partir de lo obtenido en PET
 - Separación de las distintas ROI obtenidas

- Anotación de las ROI
 - Asignación de posición relativa de cada ROI en el cuerpo
 - Caracterización de cada ROI en base a los resultados obtenidos con cada módulo
 - Integración con el módulo de anotación de notas clínicas
- Anotación de notas clínicas
- Análisis de las notas clínicas del paciente/s
 - Clasificación de las partes del texto analizadas en base a algoritmos de aprendizaje entrenados con notas anteriores
 - Continuación de dicho entrenamiento para perfeccionar el modelo
 - Estructuración de los datos en una base de datos

En la siguiente (figura 2.1) se puede ver el esquema anterior de forma gráfica:

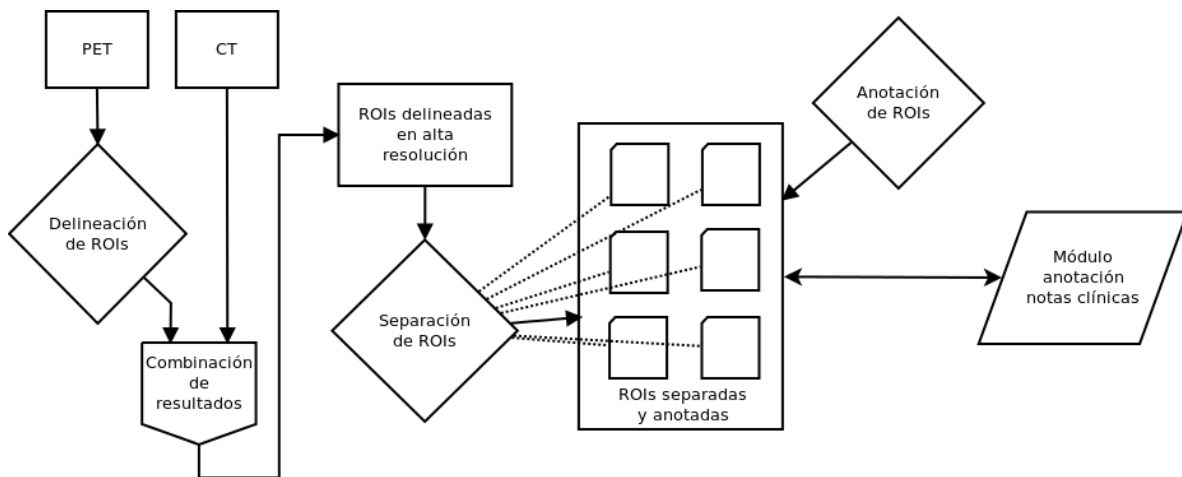


Figura 2.1: Esquema de la visión general del proyecto

Así esperamos que los especialistas tengan muchas menos dificultades y poder agilizar su trabajo sin necesidad de que tengan que actualizar su equipo ya que estarían utilizando tan sólo nuevo software. Los aparatos y dispositivos que llevan utilizando hasta ahora serán los mismos.

2.2. Objetivos específicos de este proyecto de Fin de Grado

Los objetivos generales del proyecto global han sido señalados en la sección anterior. Este proyecto de Fin de Grado se centra en la delimitación de las regiones de interés (ROI) en las imágenes PET. Ésta es la parte referida en la vista general como delineación a baja resolución.

La parte correspondiente al tratamiento de las imágenes CT no será incluido en este proyecto pero, sin olvidar que los resultados obtenidos deben poder usarse en éstas y en el posterior proceso de combinación o fusión de resultados PET con CT.

Entonces, el objetivo principal de este proyecto es la obtención de los volúmenes cancerígenos correspondientes en las imágenes PET de la manera más precisa dentro de las posibilidades.

Para conseguir este objetivo principal, se definen los siguientes objetivos específicos:

- O1:** Estudio de las técnicas utilizadas hasta la fecha para escoger los métodos que, en el tiempo dado, sean los más indicados para resolver el problema.
- O2:** Definición de la arquitectura del módulo.
- O3:** Implementación y desarrollo del submódulo de delineación de ROIs.
- O4:** Validación de resultados.
- O5:** Comparación de los resultados con los diversos métodos escogidos.
- O6:** Documentación del módulo.

Aparte de estos objetivos específicos, se pretende, si fuera posible la separación de las diversas ROI en función de los datos de posicionamiento relativo al cuerpo obtenidos mediante otro módulo del proyecto general (objetivo **O7**). Este objetivo queda entonces como objetivo adicional ya que no es en lo que se centra realmente este proyecto.

El objetivo de integración del submódulo con los demás módulos que componen el proyecto general a nivel de departamento no será incluido en la lista de objetivos y también quedará como objetivo adicional ya que los demás submódulos todavía no han llegado a ser desarrollados. Esto entonces incluye, la integración del submódulo con la delineación de alta resolución en CT.

Se considera de mayor importancia que este trabajo se centre en revisar más técnicas existentes y evaluar la corrección de los resultados en lugar de cumplir el anterior objetivo adicional u otros posibles objetivos adicionales.

Capítulo 3

Estado del arte

En este capítulo se presentará con mayor detalle lo que se ha revisado en la literatura existente en lo referente al tema de la delimitación de regiones de interés (ROI) en las imágenes de tipo PET.

Posteriormente, se pasará a describir en detalle aquellas técnicas que se han revisado para realizar la tarea de delineación de PET así como los problemas encontrados en la literatura. Las técnicas escogidas para este proyecto y la descripción de la metodología usada para realizar las pruebas no se describirán en detalle hasta el capítulo 4.

3.1. Conceptos previos

3.1.1. Segmentación de imágenes

El primer concepto a revisar es la *delimitación* de estas ROI que estamos hablando. Hablaremos indistintamente a lo largo de este documento de *delimitación*, *delineación* y de *definición* de las ROIs ya que en el contexto en el que estamos trabajando, sus significados los consideramos intercambiables.

La delimitación de regiones cancerígenas en las imágenes PET es un caso particular de los problemas conocidos como *segmentación de imágenes*. Se entiende por segmentación el proceso de dividir los píxels, vóxeles o, en general diremos, unidades de una

imagen en una serie de conjuntos disjuntos¹. A estos conjuntos disjuntos se les conoce como *segmentos* en el área de imágenes. Podríamos resumirlo matemáticamente en que segmentar una imagen de cualquier tipo consiste en crear una partición sobre el conjunto de unidades que componen la imagen donde cada parte se corresponde con un segmento.

La unidad en las imágenes sirve para relacionar un fragmento de la imagen único con un dato de nuestro interés. Esto es lo que se conoce en el día a día como *píxel* si la imagen es de dos dimensiones o *vóxel* si la imagen es de tres dimensiones. En medicina se suele tratar con ambos tipos de imágenes así que cuando nos refiramos a la unidad en las imágenes hablaremos directamente de vóxeles o píxeles dependiendo de las dimensiones de la misma.

El dato de interés que se relaciona con la posición del píxel o del vóxel dependerá de la información que se quiera obtener de la imagen y, por supuesto, de la información que proporcione la imagen. Normalmente, son interesantes datos de color o intensidad del píxel o vóxel en una dimensión. En el caso de ser de interés el color, se suelen usar 3 canales o incluso 4 canales de color, como los conocidos canales RGB (red, green, blue) y RGBA (red, green, blue, alpha). Aunque en las aplicaciones médicas es más común usar tan sólo un canal de intensidad cuyo rango de valores varía entre unos tipos de imágenes y otras. Además depende también del tipo de escáner utilizado. Se suele llamar *profundidad* al número de bits que cada píxel puede almacenar, esto afecta a la cantidad de valores distintos que se pueden asociar a un píxel o vóxel en una imagen.

En general, la segmentación de imágenes es útil para extraer características de la imagen, para hacer medidas en regiones concretas y para mejorar la visualización de las imágenes. En el área de medicina, es interesante distinguir tipos de tejidos distintos como huesos, músculos y venas, y en otras, interesa distinguir tejidos cancerígenos, deformidades de tejidos o lesiones de esclerosis múltiple [1].

Como se puede observar en todos estos casos, el problema se ajusta al problema de segmentación de imágenes ya que, en general, de lo que se está tratando es de distinguir las diferentes regiones de la imagen y clasificarlas acorde con lo que representan en la realidad. La clasificación entonces es importante llevarla a cabo correctamente, aunque la gravedad de las consecuencias por obtener resultados erróneos depende del ámbito en que se esté trabajando.

¹Recordar que una serie de conjuntos disjuntos cumple que cada uno de los elementos o unidades que lo compone pertenece sólo a ese conjunto y a ninguno más de los otros.

3.1.2. Proceso de segmentación

Para poder llevar a cabo el proceso de segmentación de imágenes es necesario saber qué queremos sacar de la imagen o imágenes que queremos segmentar y además, conocer la naturaleza de las mismas puesto que dependiendo de la forma, estructura y distribución de la información que proporcione la imagen, hará falta resolver la solución del problema siguiendo una estrategia u otra.

Entonces, lo primero que se debe identificar antes del proceso de segmentación es el objetivo que tenemos, es decir, qué información queremos extraer. Una vez identificado, debemos entender cómo se presenta la información en la imagen para elegir una *técnica de segmentación* y posteriormente, probarla sobre la imagen o conjunto de imágenes para ver los resultados obtenidos.

En algunas ocasiones, será necesario el uso de múltiples técnicas durante el proceso para conseguir la segmentación deseada [1]. El uso de múltiples técnicas se puede llevar a cabo tanto en paralelo como en secuencia o incluso combinar en paralelo y en secuencia.

Habrán casos en los que las técnicas en secuencia utilizadas no se usen necesariamente para la segmentación en sí, sino que más bien, se utilizan para adaptar ciertas condiciones de la imagen que no nos son favorables en el posterior proceso de segmentación. Estas técnicas previas a la segmentación se conocen en general en el campo de tratamiento de imágenes como *técnicas de preprocesamiento*. Estas técnicas suelen usarse para reducir la cantidad de información que proporcionan estas imágenes ya que a menudo, contienen más información de la deseada y dificulta la búsqueda de patrones y, por tanto, la segmentación de las imágenes. Por ejemplo, el ruido que suele haber en muchas imágenes dificulta la segmentación de las mismas y hacer uso de ciertas técnicas de preprocesamiento tal y como puede ser el conocido *filtro gaussiano* pueden ayudar a reducir el ruido de la imagen.

Una vez que ya tenemos la imagen lista para que sea segmentada, podemos decidirnos por una o varias técnicas de segmentación y continuar el proceso. Las diversas técnicas de segmentación existentes y más populares en el ámbito de la medicina serán descritas en la sección 3.2 y se darán algunos ejemplos de sus aplicaciones, de nuevo, en el área de imágenes médicas.

3.1.3. Características de las imágenes PET y CT

Ahora vamos a describir en más detalle las características de las imágenes CT y sobre todo PET así como su uso en el tratamiento de pacientes con cáncer y las dificultades a las que se enfrentan los especialistas cuando tienen que delinear tumores cancerígenos usando estos tipos de imágenes.

Como vimos en la introducción, las mayores diferencias presentes en las imágenes PET y CT se daban en su sensibilidad y su resolución. Esto afecta severamente a la nitidez de las imágenes y la facilidad para distinguir las regiones de interés.

Una característica común tanto en las imágenes PET como las imágenes CT es que ambas son tridimensionales y, en ambas pueden abarcar prácticamente todo el cuerpo del paciente, aunque el escáner se hace normalmente tan sólo del tórax y del abdomen de los pacientes. Para hablar de las coordenadas que se manejan en estas imágenes usaremos los ejes X e Y para referirnos a las coordenadas de cada una de las imágenes bidimensionales obtenidas durante el proceso. Cada una de estas imágenes se conoce como *slice* y recorren el eje que llamaremos Z cada vez que sea necesario. Cada uno de los *slices* del paciente se corresponde con un “sección” del mismo en lo que se conoce como plano *transversal* del cuerpo.

En oncología, las imágenes CT no son suficiente para poder delimitar o delinear los volúmenes de los tumores cancerígenos. El problema que se da con las imágenes CT es que no hay marcadores equivalentes a los de PET que puedan hacer que los tumores malignos queden correctamente definidos. Sin embargo, al ser imágenes de relativamente alta resolución, los resultados de la delineación serían bastante más precisos aunque también sería mucho más laborioso. Pero esto no se puede aplicar a los volúmenes cancerígenos ya que a la vista de un especialista los tejidos son prácticamente indistinguibles en muchos casos debido a lo que se acaba de comentar sobre los tejidos blandos.



Figura 3.1: Ejemplo de imagen CT de tórax

En otros tipos de tejidos, la delinea-

ción es una tarea relativamente sencilla de realizar por un especialista. Por ejemplo, los tejidos óseos son muy fáciles de distinguir frente a los tejidos blandos que pueda haber alrededor (suele haber mucho contraste). No siempre es sencillo automatizar estas tareas ya que en este tipo de imágenes nos podemos encontrar con la presencia de artefactos o simplemente, la dificultad conocida que tienen los algoritmos de reconocimiento de patrones visuales en general.

Dentro del campo de oncología, se usan entonces las imágenes PET ya que son imágenes en las que el marcador ^{18}F -FDG se puede distinguir visual y analíticamente de las demás zonas debido al alto contraste. No obstante, como indicamos previamente, el problema de estas imágenes es la baja resolución y la presencia de ruido. Esto significa que los segmentos que obtengamos tras hacer el proceso de delineación en estas imágenes no serán tan homogéneos como teóricamente deberían ser, con lo cual, es sencillo cometer errores en la delineación automática de este tipo de imágenes.

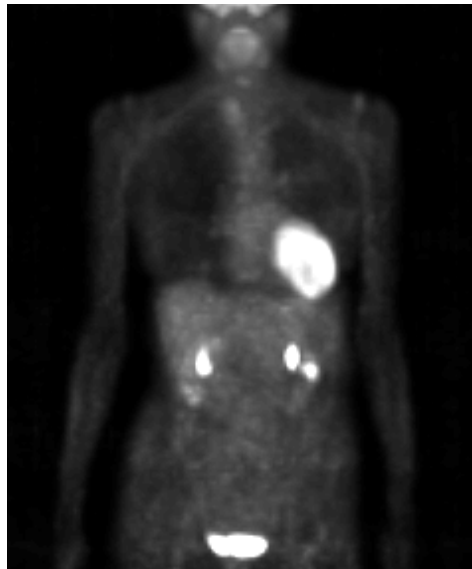


Figura 3.2: Ejemplo de imagen PET de paciente con cáncer de pulmón

En la figura 3.2 se puede observar un ejemplo de imagen PET donde los tonos de color blanco se corresponden con valores de actividad metabólica superiores y los de color negro se corresponden con valores inferiores. Como sabemos, la alta intensidad está directamente relacionada con que el escáner ha recibido muchos positrones de esa zona, lo que significa que hay mucho ^{18}F -FDG concentrado en dichos puntos y ello implica que hay mayor probabilidad de que exista un cáncer de algún tipo. No obstante, tan sólo indica que hay un mayor metabolismo en la zona y esto suele estar relacionado con el cáncer, por eso se ha dicho “probabilidad” en la línea anterior. En esta imagen, el cáncer está localizado en el pulmón izquierdo del paciente.

No saber si se trata exactamente de un tumor cancerígeno puede dar muchos problemas a la hora de averiguar qué zonas se deben considerar como falsos positivos. Sin embargo, se sabe de antemano que zonas como el cerebro, los riñones, la próstata y la vejiga, tienen una mayor concentración de marcador (^{18}F -FDG) [6]. El cerebro tiene mayor concentración puesto que siempre tiene mayor metabolismo. Pero esto no es lo que sucede con los riñones y la vejiga, en estos casos, la mayor concentración de marcador se debe a que al realizar el cuerpo la función de excreción, almacena grandes cantidades de ^{18}F puesto que la glucosa ya ha sido consumida durante el proceso de metabolismo. Por tanto, cuanto más tiempo pase desde que se inyecta el marcador en el paciente, menos disperso está el marcador ya que se da tiempo a que se concentre en las zonas de metabolismo más alto; pero si la cantidad de tiempo transcurrido es demasiado alta, es posible que la mayor parte de marcador quede en la vejiga. En la figura 3.2 se pueden observar todas estas zonas que se acaban de comentar.

En casi todas las técnicas conocidas, no se trabaja directamente con las intensidades de las imágenes obtenidas mediante los escáneres PET. Normalmente, se hace una conversión a otra unidad que ya ha sido introducida anteriormente y que se llama *standard uptake value* (SUV). El SUV es una medida también de la actividad metabólica directamente relacionada con la cantidad de actividad metabólica en cada vóxel en función del tiempo, del peso del paciente y de la cantidad de dosis inyectada [6, 24]. No tener en cuenta estos factores puede acarrear problemas ya que dependiendo del caso, los valores obtenidos pueden variar notablemente entre pacientes y se hace más complicada la comparación de los valores en los mismos. Por tanto, el SUV apareció como un intento de unificar de alguna manera los valores obtenidos en todos los pacientes.

El SUV es concretamente una medida de corrección sobre la vida media del marcador por cada unidad de masa corporal y por cada dosis administrada en función del peso del paciente. En otros términos, el SUV se puede expresar como indica la fórmula 3.1 extraída del artículo [24].

$$SUV = \frac{cm}{di/w} \cdot \frac{1}{d} \quad (3.1)$$

donde cm es la concentración media de marcador en la ROI y se mide en MBq/g donde g son gramos y donde MBq se entiende por millones de Bq (*becquerel*) que es una medida utilizada en el sistema internacional (SI) para medir la radiactividad; di es la cantidad de dosis inyectada medida en MBq , el peso del paciente w medido en gramos y el factor de tiempo de descomposición del marcador. Por tanto, con el SUV, se tiene en cuenta el tiempo que ha pasado desde el comienzo de las pruebas en el escáner. De hecho, incluso se tiene en cuenta el tiempo transcurrido entre un *slice* y otro.

Es útil también conocer el SUV máximo del tumor cancerígeno en el paciente, que en muchas ocasiones se usa como referencia para comprobar el estado de avance en el

que se encuentra el cáncer y para ayudar a preveer la evolución de la enfermedad [6].

El problema que tiene el uso del SUV es que la fiabilidad del mismo depende de una serie de factores técnicos y de los pacientes [24, 11]. Factores como el tiempo que transcurre entre el momento de inyección y el momento en el que comienzan las pruebas o como el grado de infiltración del marcador en la zona de inyección pueden afectar la corrección y fiabilidad del SUV, no siendo siempre la medida más indicada [24]. Sin embargo, esta medida es la que se ha aceptado comunmente como la más acertada y es la que se suele utilizar en los hospitales actualmente.

Para poder evitar los problemas que tienen las imágenes PET y CT y poder aprovechar los datos que proporcionan ambos tipos de imágenes, se hace uso del híbrido PET-CT que ya fue introducido al comienzo de este documento. Con las imágenes PET-CT estaríamos aprovechando la resolución espacial de las imágenes CT y la sensibilidad de las imágenes PET.

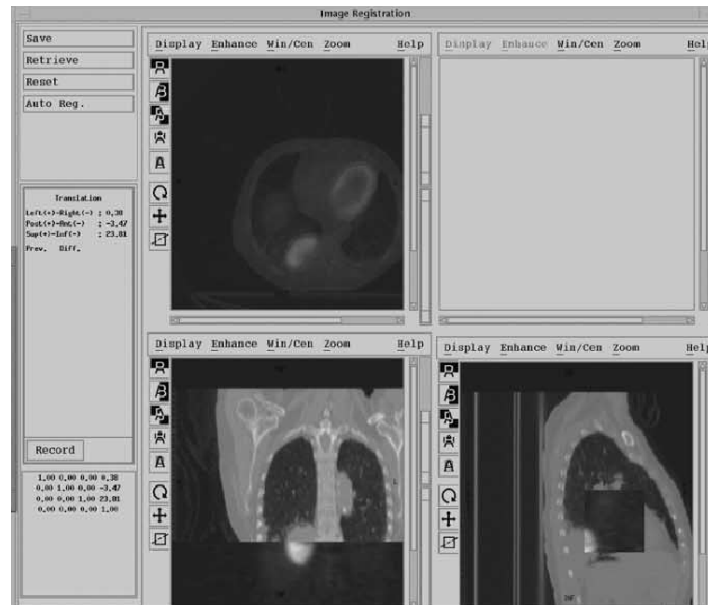


Figura 3.3: Aplicación de visualización PET-CT [5]

Las imágenes PET-CT se suelen visualizar usando aplicaciones software (ver figura 3.3) que permiten dividir la interfaz de usuario en dos ventanas simultáneamente de forma que en una ventana se muestran los resultados obtenidos por el escáner CT y los obtenidos por el escáner PET. La imagen 3.4 es un ejemplo de cómo se representaría el híbrido PET-CT en tres dimensiones.

Estos escáneres están ya integrados para poder analizar a los pacientes y obtener

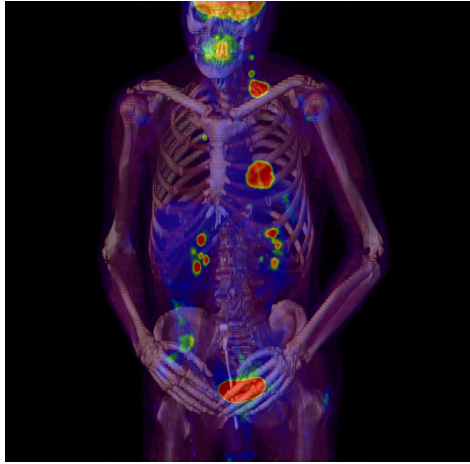


Figura 3.4: Ejemplo de híbrido PET-CT 3D

las imágenes PET y CT al mismo tiempo. Esto es importante ya que si se tuviesen que obtener en dos sesiones de pruebas distintas, se correría el riesgo de que el paciente no esté alineado de la misma manera en ambas pruebas y entonces, la información es más difícil de tratar, sobre todo si se hace uso de técnicas automáticas.

La combinación de imágenes PET y CT se hace hoy en día para evitar problemas que se dan al delinear volúmenes tanto si se hace sólo en PET como si se hace sólo en CT. Los volúmenes delineados en PET garantizan que si el tumor cancerígeno ha sido captado por el escáner, es mucho menos probable que el operador entrenado en delimitación de contornos de tumores cancerígenos en imágenes PET cometa el error de dejar el volumen sin delinear. Pero por otro lado, la baja resolución de las imágenes hace que la definición de los contornos no sea del todo acertada. Delimitar correctamente los volúmenes es importante porque en casos como los carcinomas pulmonares no microcíticos (*non-small-cell lung cancer* o NSCLC) es normal que se requiera de radioterapia ya que la quimioterapia no suele servir. Y durante el tratamiento mediante radioterapia, es imprescindible conocer el volumen cancerígeno real con la mayor exactitud posible puesto que la radioterapia es de tipo ionizante y si se cometen errores, se pueden afectar a zonas sanas pudiendo tener consecuencias graves.

En cuanto a las imágenes CT, el problema reside en que hay ocasiones en las que no se detectan ciertas zonas afectadas porque a la vista de los especialistas no lo parecen. Sin embargo, se puede conseguir una delineación del tumor mucho más exacta y por consiguiente, la radioterapia tendrá menos efectos nocivos.

En dos estudios [5, 7] de NSCLC se demuestra que estos errores se cometen con frecuencia al delinear sólo CT y que, consecuentemente, hacer uso de PET-CT mejora

los resultados obtenidos. En la siguiente figura (3.5) se muestra uno de los casos presentados en dicho estudio en los que el operador no detectó una de las zonas en CT, pero sí se detectaron en PET y, por tanto, fueron añadidas posteriormente en PET-CT obteniendo un volumen considerablemente mayor que el anterior.

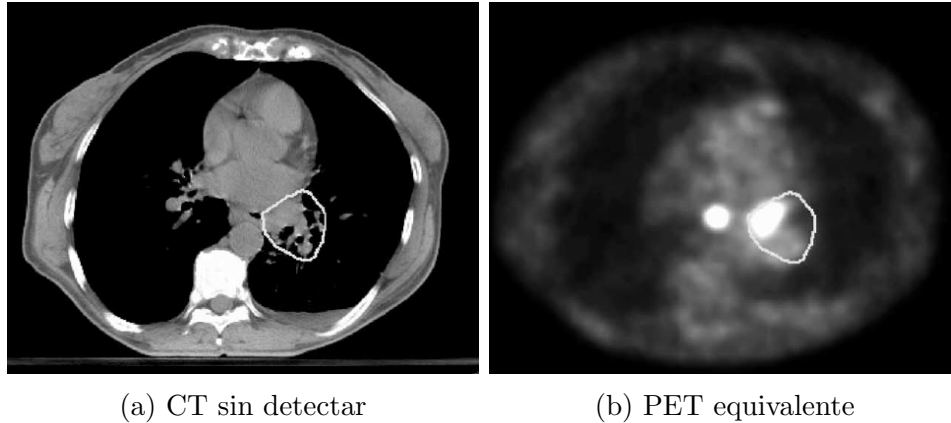


Figura 3.5: Ejemplo de región sin detectar en CT pero visible en PET [5]

Con toda esta introducción de conceptos y técnicas más o menos generales, concluye esta sección y se pasa a revisar las técnicas encontradas en la literatura centrándonos en el área de oncología y las imágenes CT y, sobre todo, PET ya que éste es nuestra principal objetivo.

3.2. Técnicas principales de segmentación

En la sección 3.1.2 hemos visto que una vez que tenemos la imagen adecuada y hemos observado su naturaleza, podemos empezar a buscar una técnica o conjunto de técnicas que nos sean más adecuadas para la segmentación definitiva de la imagen.

Para poder determinar la clasificación de cada píxel o vóxel en un segmento u otro, se suelen aproximar las soluciones usando técnicas de reconocimiento de patrones común en los campos de aprendizaje de máquinas (*machine learning*) y de minería de datos (*data mining*). Estas técnicas se adaptan al problema ya que, al fin y al cabo, lo que se intenta obtener de estas imágenes es un conjunto de segmentos donde exista una homogeneidad en algunas características entre todos los elementos que componen ese segmento. Esta homogeneidad debe ser evidentemente menos notable o incluso inexistente entre elementos que pertenecen a distintos segmentos ya que por esa razón han sido clasificados así [1].

Para facilitar la elección de las técnicas de segmentación, es muy conveniente saber cómo se comportan estas técnicas en el sentido de saber cómo trabajan con la imagen de entrada y que hacen en el resto del procedimiento o algoritmo. Además, hay técnicas que requieren de cierta información adicional de entrada en mayor o menos medida y, dependiendo del caso, puede que esto nos interese o no. Entonces, al existir muchas técnicas y muchos derivados de éstas, se va a hacer una clasificación que pueda ayudar a distinguir las técnicas por sus características comunes frente a ciertos criterios de clasificación.

Clasificaremos entonces usando los siguientes criterios de clasificación [1]:

Grado de interacción - según el grado de interacción que requiera la técnica con respecto al usuario. Aunque como se puede comprender, es complicado decidir hasta qué nivel una técnica es completamente automática o si más bien debería ser vista como semiautomática. En caso de duda, se dará un razonamiento de por qué la clasificamos de una manera o de otra.

- Manuales: grado de interacción máximo.
- Semiautomáticas: requieren ciertos parámetros que se consideran excesivos como para ser clasificadas en el siguiente grupo.
- Automáticas: no requieren interacción en absoluto o bien se requieren tan sólo unos pocos parámetros de menor importancia para la clasificación.

Localización - en este caso estamos refiriéndonos a la localización o posición de la información que la técnica extrae de la imagen.

- Globales: que usan toda la información de la imagen al mismo tiempo.
- Locales: basadas en regiones independientes de la imagen pudiendo ser incluso un solo píxel o vóxel.

Nivel de abstracción - al igual que sucedería con el nivel de los lenguajes de programación en función a lo cercanos que son al *hardware* de la máquina, en nuestro caso nos referimos al nivel de manipulación de la información que se maneja.

- Delineación manual: manipulación al nivel más bajo posible, no hay ninguna “capa” de abstracción entre medias.
- Segmentación de bajo nivel: técnicas que trabajan con los datos de la imagen sin añadir significado como la umbralización y las técnicas de crecimiento de regiones.

- Segmentación basada en modelos: técnicas que tratan de relacionar datos usando modelos y, por tanto, aumentando el nivel de abstracción y de significado a la información extraída. Algunos ejemplos de técnicas de este estilo tales como técnicas multimodales y técnicas de programación dinámica.

Rama de la ciencia - podríamos decir que ésta es la clasificación clásica en el sentido de que se utilizan directamente grupos de técnicas conocidos en otras ramas de la informática.

- Clásicas: aquéllas técnicas que se diseñaron específicamente para el análisis de imágenes en general tales como la umbralización, las basadas en contornos y las basadas en regiones
- Basadas en estadísticas.
- Basadas en lógica borrosa (*fuzzy logic*).
- Basadas en redes neuronales.

Se hará ahora una breve introducción a las técnicas más populares en segmentación de imágenes médicas sin necesidad de seguir ningún criterio de clasificación específico:

Técnicas de umbralización: estas técnicas son las más comunes dentro de las técnicas globales [1]. Las técnicas de umbralización consisten en establecer un *umbral* sobre los valores (intensidad) que pueden tomar los píxeles o vóxeles de la imagen. De esta manera, se distinguen al menos dos clases en función del umbral seleccionado. Estas dos clases son, la clase cuyos píxeles o vóxeles tienen dicho valor por encima del umbral o igual y la clase de aquéllos que están por debajo de dicho umbral. Aunque esta definición sólo sirve para los casos en los que se usa un solo umbral global, es decir, un único umbral aplicado a toda la imagen.

En general, podemos definir estas técnicas de la siguiente manera asumiendo que se trabaja sobre un solo canal de intensidad de las imágenes. Las técnicas de umbralización se basan en clasificar todos los píxeles, vóxeles o la unidad en la dimensión que sea siguiendo la fórmula 3.2.

$$g(\bar{x}) = \begin{cases} 1 & \text{si } f(\bar{x}) \geq u \\ 0 & \text{si } f(\bar{x}) < u \end{cases} \quad (3.2)$$

siendo $f(\bar{x})$ la imagen original y $g(\bar{x})$ la imagen binaria resultante. Además, en general, llamaremos *objeto* (o *foreground*) al segmento cuyos píxeles o vóxeles sean clasificados como 1 y *fondo* (o *background*) a aquéllos que se hayan clasificado como 0.

La fórmula anterior sólo nos sirve para dar un solo umbral global, sin embargo, existen técnicas que dan umbrales diferentes a cada zona o incluso a cada píxel o vóxel de la imagen. En este caso, bastaría con sustituir u en la fórmula por $u_{h(\bar{x})}$ donde $h(\bar{x})$ sería una función que nos devuelve el índice del umbral en la lista de umbrales para cada píxel, vóxel o en la dimensión que sea con coordenadas (\bar{x}) en la imagen. Es decir, que se asume además la existencia de una lista de umbrales para llevar a cabo este proceso de segmentación descrito mediante la fórmula.

Por otro lado, hay técnicas de umbralización que se basan en establecer varios umbrales independientes en partes disjuntas de la imagen (*umbral local*) o que incluso establecen múltiples umbrales para obtener varios segmentos de objeto distintos (siempre se entiende que sólo existe un segmento de fondo) en cuyo caso diremos que se trata de *multiumbralización* la cual será descrita con mayor detalle en la sección 3.4.2.

En conclusión, podemos comprobar que las técnicas de umbralización se centran en buscar diferencias entre intensidades de píxeles o vóxeles para establecer el umbral. Aunque, en realidad, el resultado sigue siendo una agrupación de píxeles o vóxeles con características similares.

Se mostrarán ejemplos de imágenes obtenidas mediante este tipo de técnica más adelante, concretamente, en la sección 3.4.1.

Técnicas de crecimiento de regiones: las técnicas de crecimiento de regiones en lugar de buscar la diferencia entre píxeles o vóxeles, tratan de agrupar aquéllos que tengan ciertas similitudes en alguna característica. Este tipo de técnica comienza haciendo uso de una semilla (*seed* en inglés) y un posterior proceso de crecimiento de la región a partir de dicha semilla. La semilla deberá ser un píxel o vóxel que se sabe que pertenece o predice que pertenece a dicha región. La semilla puede ser introducida tanto por un operario en el programa como elegida mediante otro algoritmo [1]. Por tanto, el procedimiento podemos dividirlo en selección de la semilla, crecimiento de la región y determinación del fin del proceso de crecimiento ya que no puede crecer indefinidamente.

El criterio utilizado para decidir cuál es el siguiente píxel o vóxel a introducir en la región se conoce como *criterio de homogeneidad* o *criterio de uniformidad* [1]. Este criterio normalmente establece píxel a píxel o vóxel a vóxel cuáles son los siguientes que van a ser incluidos o clasificados como región de interés. La elección de los siguientes píxeles o vóxeles a clasificar se suele hacer por proximidad a la semilla o a los últimos que han sido clasificados como positivos. Este proceso termina cuando no se consigan incluir más píxeles o vóxeles en el conjunto una vez evaluados todos aquéllos que deban

ser evaluados en la siguiente iteración.

La ventaja de este tipo de técnica es que se extraen los píxeles y vóxeles con características similares relativas a la intensidad en cada punto, pero además, se tienen en cuenta las características espaciales en la imagen. Por esta razón, las regiones obtenidas con estos procedimientos son siempre regiones conexas, lo cual sería necesario comprobar en otros tipos de métodos si es que se requieren regiones necesariamente conexas. Por otro lado, una desventaja que presentan es que se obtienen distintos resultados dependiendo de qué semilla se elija. Pero no sólo depende de la semilla, también depende enormemente del criterio de homogeneidad que se escoja.

Un ejemplo de crecimiento de regiones es el uso de *hill climbing* en la delineación de microcalcificaciones en mamogramas. Para usar esta técnica, se comienza escogiendo el máximo local que hay en dicha microcalcificación y se usa como semilla. Desde este punto se dibujan de forma radial una serie de líneas con el mismo ángulo entre todas ellas. Estas líneas determinan las direcciones en las que se va a realizar el *hill climbing*. El *hill climbing* consiste en avanzar píxel a píxel en la dirección escogida hasta alcanzar un máximo en una función determinada [1]. En este caso, la función representa la inclinación (en el sentido de gradiente) que hay entre unas intensidades y otras. El punto de inclinación máxima localmente es el punto escogido para cada línea trazada.

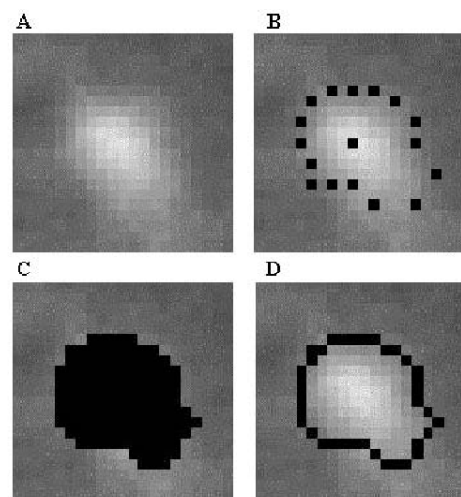


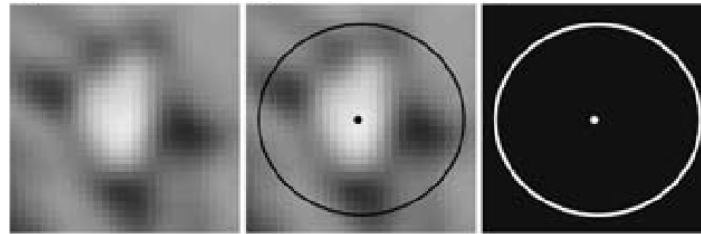
Figura 3.6: Proceso descrito de *hill climbing* [1]

Como podemos ver en la figura 3.6, se escoge el máximo local de la zona de mayor intensidad de la imagen y desde ahí se alzan 16 líneas equiespaciadas de forma radial para calcular mediante *hill climbing* los 16 puntos que hacen de límite de la región. Estos puntos se usan después como semillas para hacer crecer la región hacia dentro, es decir, en este ejemplo el criterio de homogeneidad hace que se escojan píxeles que se acerquen al máximo local tanto espacialmente como en términos de intensidad.

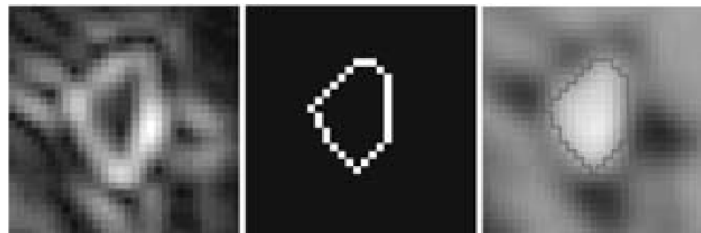
Este tipo de técnica se ha utilizado también en otros casos como puede ser la segmentación de ventrículos en imágenes cardíacas, la extracción de venas en angiografías

y la extracción de la superficie del cerebro en otras áreas de imágenes médicas [1].

Algoritmo de Watershed: la segmentación mediante el algoritmo de Watershed está basada en regiones y utiliza la morfología de la imagen. De nuevo, se requiere una semilla interior a la futura región, pero además, se debe marcar también la zona a la que pertenece el fondo conocido de la imagen. Para ello, se puede hacer esto, de nuevo, mediante un operador, o bien, mediante un algoritmo que marque ambas cosas automáticamente usando conocimientos previos² sobre la imagen [1].



(a) Elección del punto interior y del fondo



(b) Algoritmo de Watershed tras aplicar el operador Sobel

Figura 3.7: Ejemplo de algoritmo de Watershed descrito [1]

Este algoritmo funciona intuitivamente de la siguiente forma (figura 3.7). Si asociamos la intensidad de los píxeles de la imagen con la altitud de un terreno, la imagen tridimensional resultante sería la de un terreno con picos, montañas y valles a distintas altitudes. Si pudiésemos sumergir dicho terreno en agua de forma que el eje de altitud del terreno fuese paralelo al eje de profundidad en una bañera, este algoritmo seleccionaría los puntos en los que el agua comenzaría a invadir otros valles distintos del valle externo. Es decir, básicamente, este algoritmo es capaz de distinguir los bordes de las “piscinas” que se formarían en el terreno o, lo que es lo mismo, sería capaz de marcar cordilleras [1]. Análogamente, si simulásemos la caída de gotas de agua en cada uno de los puntos del terreno, podríamos decir que el algoritmo es capaz de detectar qué puntos tienen una vertiente que da a un valle o a otro. Los puntos que lleven al valle donde

²Aquellos que se conocen debido a la propia naturaleza del tipo de imagen.

está la semilla interior son puntos pertenecientes a la “piscina” que nos interesa, los demás pertenecen al fondo.

No todas las imágenes tienen esta característica de que nos interese marcar sus cordilleras, de hecho, es más usual querer delimitar tan sólo la región de mayor intensidad de estas imágenes. Por tanto, lo que se puede hacer es convertir dichas imágenes usando el operador Sobel³ para obtener una imagen donde la intensidad de los píxeles es proporcional a la diferencia de intensidades entre píxeles vecinos, que en este caso, sí coinciden con el límite de dicha región [1].

Esta técnica se utiliza en numerosas aplicaciones en imágenes de dos dimensiones. Por ejemplo, es útil para segmentar imágenes microscópicas de la retina. También se ha llegado a implementar una versión tridimensional para analizar imágenes cardíacas volumétricas [1].

Técnicas basadas en contornos: estas técnicas consisten en usar los contornos detectados en las imágenes para hacer una posterior segmentación de la misma. Al igual que en el caso anterior, lo que se hace es hallar los contornos usando un operador, normalmente basado en el *gradiente*, tal y como sucede con el operador Sobel (ver figura 3.8 B). Por tanto, la estrategia a seguir es la misma que la de antes, primero se obtienen los bordes de la imagen usando cierto operador y, después, se determina qué bordes nos interesan [1].

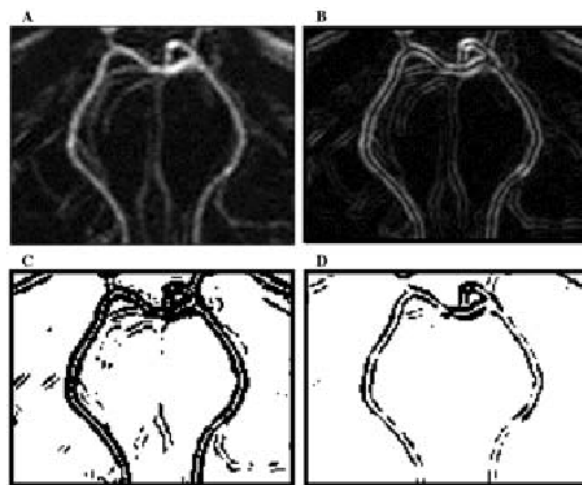


Figura 3.8: (A) imagen original, (B) tras aplicar el operador Sobel, (C y D) tras umbralizar con distintos umbrales [1]

³Este operador está basado en el gradiente.

En este caso, se suele hacer uso de técnicas de umbralización sobre la imagen de gradientes para determinar cuáles son los bordes que nos interesa clasificar como bordes (ver figura 3.8 C). Pero en muchas ocasiones, estos bordes no formarán recintos cerrados, lo cual suele ser un requisito imprescindible en este tipo de segmentación. Para poder cerrar los recintos tras haber realizado el proceso anterior, se suele hacer uso de técnicas de postprocesado que cierran dichos recintos. Esto se puede hacer de forma automática, aunque puede resultar ser una situación bastante ambigua. En este caso, será necesario hacer uso de métodos semiautomáticos o completamente manuales [1]. Otra posibilidad es la de usar el operador Laplaciano en lugar de otros operadores, que incluye la información que proporcionan las segundas derivadas.

Estas técnicas suelen presentar problemas en casos de ruido en las imágenes ya que el ruido puede alterar significativamente los resultados y generar contornos que no nos interesen [1].

Diversas versiones de estos métodos se han utilizado para analizar imágenes médicas. Se ha utilizado para detectar las cámaras ventriculares en imágenes de resonancia magnética del corazón, para segmentar imágenes de resonancia magnética del cerebro y para delinear venas [1].

Técnicas multimodales: las técnicas multimodales se basan en utilizar la información que proporcionan varias imágenes de distinta naturaleza al mismo tiempo. La variedad de la información que proporcionan dichas imágenes depende del tipo de imagen que se trate. Podemos estar combinando características de imágenes que provienen de distintos tipos de escáner (MRI, PET, CT...) o bien podemos estar combinando imágenes del mismo tipo de escáner pero adquiridas en distintos momentos, dándonos información más completa sobre la evolución de la enfermedad en el paciente.

Combinando múltiples técnicas: en el caso de usar varias técnicas lo que tenemos es una imagen que, en lugar de tener un solo canal de intensidad, tiene varios canales de intensidad (como sucede con las imágenes a color). Para poder sacar la información de estas imágenes correctamente en múltiples imágenes simultáneamente, se suele hacer uso de técnicas conocidas como puede ser *k-nearest neighbors* (kNN), *fuzzy c-means*, *k-means* o las *redes de neuronas artificiales* (RNA). Además, es importante que estas imágenes tengan poco ruido porque deben corresponderse las unas con las otras. El ruido y otros factores puede hacer que relacionar imágenes sea mucho más complicado [1].

Este tipo de técnica ha resultado ser muy eficiente por ejemplo en el análisis del

cerebro en dos versiones distintas de MRI para poder contrastar los tejidos que lo componen [1].

Adquisición en función del tiempo: las técnicas más conocidas para analizar en función del tiempo se conocen como *técnicas de análisis paramétrico*. En estas técnicas la intensidad de cada elemento se contrasta con el tiempo generando unas gráficas. Se pueden elegir una serie de parámetros como la intensidad máxima y mínima que se utilizarán para obtener unas imágenes resultantes y poder hacer el estudio consecuente sobre ellas en función de un solo parámetro. Cada gráfica se analiza asumiendo que las curvas tienen características similares en el tiempo, es decir, que siguen el mismo patrón [1].

Este tipo de técnica se ha utilizado por ejemplo en segmentación tridimensional de lesiones de esclerosis múltiple en imágenes MRI obtenidas a lo largo del tiempo. Después, se escogen parámetros como por ejemplo, variación de la intensidad en las lesiones conocidas, momento de aparición o momento de desaparición de las mismas. Con esto, se obtienen unas imágenes en las que se pueden observar resaltadas otras lesiones que siguen los patrones establecidos por los parámetros elegidos.

Con esto, se han introducido en menor detalle las técnicas más populares en el mundo del análisis de imágenes médicas, se han proporcionado ciertos ejemplos sobre cómo se aplica determinada versión de la técnica y en qué ámbitos reales se ha probado. Además, se han dado ciertas nociones para entender las ventajas y desventajas que pueden tener estos tipos de técnicas en función de la situación.

3.3. Segmentación manual y semiautomática

Ahora se procede a introducir tanto la segmentación manual como la semiautomática en imágenes PET y CT en base a lo descrito en la clasificación 3.2. La información adicional que se debe proporcionar en cada técnica varía evidentemente según el contexto y según la técnica que se esté utilizando. En nuestro caso, en el campo de la oncología, las técnicas semiautomáticas suelen requerir información previa acerca de la posición de los volúmenes de cáncer. En muchos casos, basta con introducir un punto interior al volumen.

Primero se revisará la segmentación manual. La segmentación manual de los tumores cancerígenos requiere de especialistas entrenados en este campo. Por tanto, se requiere

de mucha experiencia y de mucho tiempo para llegar a delinear los volúmenes de la forma más correcta posible.

Aunque disponer de un especialista muy entrenado pueda parecer garantía de obtención de resultados suficientemente correctos, esto no es del todo cierto. Es conocido el problema de que la delineación manual no es del todo objetiva, es decir, que depende de cada especialista. Esto se le llama variabilidad entre observadores (*inter-observer variability*) y está ilustrado en la figura 3.9 [24, 12]. Además, también se sabe que existe un problema de variabilidad en la delineación manual debido a la configuración del programa que el especialista esté usando [24].

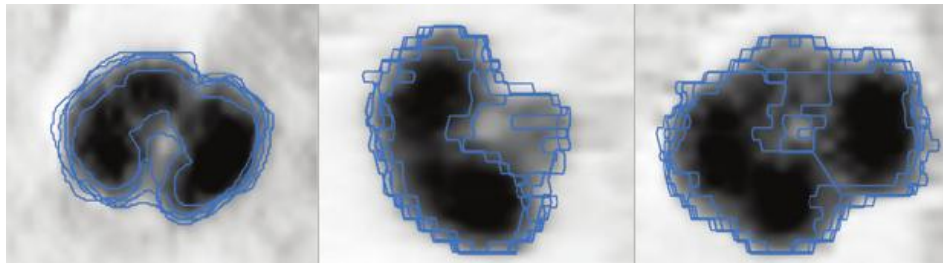


Figura 3.9: Diferencias en 3 tumores distintos entre varios observadores indicado mediante contornos azules [12]

En el caso de las imágenes CT, también existen problemas para obtener la delineación más correcta posible. De hecho, se suele sobreestimar la extensión del tumor para garantizar que la delineación no queda fuera de la lesión [8].

Para evitar estos problemas se hace uso a menudo de técnicas de crecimiento de regiones. Tal y como se explicó en la sección 3.2, estas técnicas consisten en el uso de al menos una semilla o *seed* principal la cual se usa como punto de partida para hacer crecer la región hasta el punto que se crea conveniente.

El problema que pueden llegar a presentar este tipo de métodos es que dependen mucho de la función que se use como criterio de homogeneidad (revisar 3.2) y además, depende también de cómo se elijan las semillas iniciales [8]. Esto se puede comprobar en el siguiente estudio (ver figura 3.10) donde se propone una nueva técnica de crecimiento de regiones a imágenes CT.

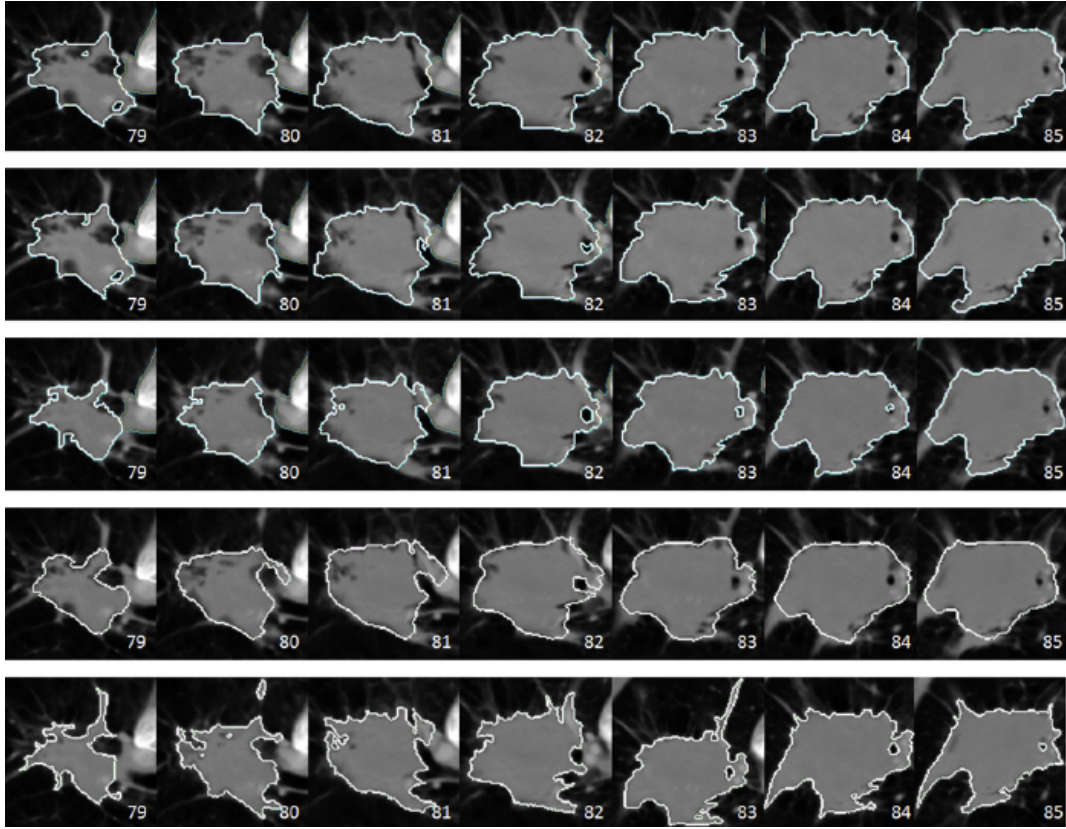


Figura 3.10: Diferentes resultados obtenidos en el mismo paciente entre distintas semillas iniciales (*slices* 79-85) [8]

3.4. Segmentación automática

Los procesos de segmentación automática o sin supervisar son los que vamos a analizar con más detalle en esta sección, sólo que esta vez nos centraremos tan sólo en PET.

La segmentación automática, como ya hemos visto, tiene la peculiaridad de no requerir la intervención en ningún momento del operario durante el proceso. Tan sólo es necesario dar unos primeros parámetros si es que la técnica que se utilice lo requiere y eso es todo. Tener en cuenta que no es sencillo clasificar entre técnicas automáticas y semiautomáticas puesto que la línea que divide a ambas definiciones es muy fina. Por ejemplo, las técnicas de crecimiento de regiones son semiautomáticas porque requieren la semilla inicial que es considerada como una ventaja importante para el procedimiento. Y por otro lado, hay técnicas que requieren ciertos ajustes para llegar a conseguir el

resultado que se pretende, por tanto, en el fondo sigue estando presente un operario. Pero como estos parámetros se pueden conseguir en realidad entrenando un modelo o a través de datos estadísticos, al final se pueden considerar como automáticos.

La segmentación automática o, en nuestro caso, la podemos llamar “delineación” automática, presenta varias ventajas especialmente frente a la segmentación manual. Principalmente, podríamos hablar de una ventaja en la falta de variabilidad frente a la que pueda haber entre varios observadores en el caso manual. Evidentemente, esto depende en gran medida en el algoritmo utilizado ya que algoritmos basados en parámetros aleatorios carecen de rigor tal y como sucedía con el crecimiento de regiones (ver 3.10). En este sentido estaríamos hablando de una mayor *reproducibilidad* de las pruebas puesto que podrían repetirse en distintas ocasiones y siempre obtener los mismos resultados. Como es lógico, esto es algo muy positivo de cara a hacer mediciones y a intentar predecir el comportamiento de las técnicas automáticas en su futuro uso.

Y además de esto, como también es evidente, el hecho de disponer de técnicas automáticas hace que no sea necesario, en nuestro caso, disponer de un operador entrenado ni de que tenga que emplear tanto tiempo en analizar las imágenes y sus respectivos volúmenes cancerígenos.

Ahora procedemos a explicar las distintas técnicas encontradas que potencialmente podrían ser útiles para ser aplicadas a lo que nos concierne que son las imágenes PET y el híbrido PET-CT. También se describirán los problemas encontrados en la revisión de la literatura en los casos que sean necesarios tales como incoherencias o falta de datos sobre los experimentos entre distintas publicaciones.

3.4.1. Técnicas de umbralización

Las técnicas de segmentación más populares en análisis de imágenes médicas ya han sido introducidas anteriormente y entre ellas estaba la umbralización (ver 3.2). Por ello nos vamos a centrar más adelante en revisar aquéllas que están disponibles en la literatura pero aplicadas a nuestro campo de estudio de las imágenes PET. Para recordar en qué consisten las técnicas de umbralización desde un punto de vista general ir al párrafo 3.2.

En esta sección distinguiremos entre dos tipos de umbralización dentro de nuestro ámbito de imágenes PET. Se distinguirá entre umbralización fija y umbralización dinámica.

Umbralización fija

Entendemos por umbralización fija a aquélla basada en un umbral constante, aunque también se suele entender como umbralización fija a umbrales fijos pero relativos al rango de valores de la imagen. Entonces, podemos distinguir entre umbrales fijos *absolutos* y umbrales fijos *relativos* respectivamente.

En la literatura se han propuesto distintos métodos de umbralización fija, aunque sólo se ha propuesto uno de umbralización fija absoluta sobre los valores del SUV (para recordar qué es el SUV, ver la descripción en la subsección 3.1). Este umbral tiene el valor **2,5** sobre el SUV y fue propuesto por Paulino et al. [16]. Este umbral se halló realizando estudios y comparativas sobre tumores malignos en pacientes de cáncer en la cabeza y el cuello.

El problema de los umbrales absolutos para distinguir entre zonas afectadas por cáncer y las zonas no afectadas es que existe cierta inhomogeneidad en los tumores, lo cual puede dificultar la búsqueda de un umbral ideal [24]. Además, puede haber muchos artefactos en las imágenes haciendo que estos métodos no sirvan [24]. De hecho, se ha probado que este método ha fallado en ciertas pruebas [24] y sigue siendo un tema de discusión en el área [24, 13]. No obstante, este método se utiliza habitualmente en los hospitales como primera aproximación al volumen ideal del tumor y posteriormente, si hay cierta inexactitud, se corrige manualmente a partir de lo hallado por esta técnica de umbralización.

Por otro lado, como hemos dicho, existen umbrales que clasificaremos como fijos, pero que son relativos a los valores de intensidad de la imagen. Estos umbrales relativos se empezaron a proponer porque según algunos estudios, el umbral ideal varía dependiendo del tamaño de la esfera (*phantom*⁴) que se use para llevar a cabo el estudio estadístico, de la relación con la actividad del fondo y de otros factores [24, 4]. Es decir, que como ya se ha dicho, queda demostrado que los umbrales fijos absolutos, no son suficientemente precisos en todos los casos.

El umbral que más se suele utilizar en la literatura es el umbral relativo constante del 42 % o, en algunas ocasiones, del 40 %. Este umbral constante se suele aplicar sobre los valores del SUV de la forma descrita en la siguiente ecuación 3.3.

$$\text{umbral} = 0,42 \times (\text{maximo} - \text{mínimo}) + \text{mínimo} \quad (3.3)$$

Es decir, que se aplica el 42 % al máximo valor de la imagen para determinar el umbral

⁴Los *phantom* son volúmenes preparados para hacer pruebas que, en este caso, se rellenan de marcador FDG.

absoluto [4]. Este estudio fue llevado a cabo por Erdi et al. en el documento [4] donde se asegura que todos los umbrales relativos constantes se encuentran siempre entre el 36 % y el 44 %. Este estudio se llevó a cabo usando *phantoms* con forma de esfera de diversos tamaños.

Umbralización dinámica

Entendemos entonces como umbralización dinámica a cualquier técnica que no esté basada en un umbral fijo ya sea relativo o absoluto. En resumen, es lo opuesto a lo que se ha descrito anteriormente.

La umbralización dinámica se puede entender también en mayor o menor medida como “fija” si contiene parámetros constantes que se han obtenido mediante estudios estadísticos, normalmente usando líneas de regresión o métodos similares. En otras ocasiones, los parámetros no tienen nada que ver con estudios hechos y tan sólo sirven para guiar al algoritmo y para conseguir unos resultados determinados. Entonces, tal y como se dijo en la clasificación 3.2, éste es uno de aquellos casos de los que no estamos seguros de cómo clasificarlos. Por el momento, diremos que la umbralización fija se da sólo en los casos descritos anteriormente, y los demás los clasificaremos directamente como dinámicas aunque haya dudas presentes.

Ahora separaremos las técnicas de umbralización dinámicas en dos grupos en función de su objetivo de diseño. Distinguiremos entonces técnicas de umbralización dinámicas clásicas y técnicas de umbralización dinámicas de PET.

Umbralización dinámica clásica

Vamos a empezar hablando de las técnicas de umbralización dinámicas clásicas que no se inventaron necesariamente con la idea de que fuesen aplicadas en un campo determinado. Es decir, en cierto sentido son técnicas de umbralización con un enfoque general. Dentro de estas técnicas clásicas, se pueden distinguir distintos tipos según la estrategia que sigan a la hora de realizar los cálculos [20]:

Basadas en la forma del histograma: se analizan las curvas (máximos locales, mínimos locales,...) de los histogramas para determinar el umbral.

Basadas en *clustering*: donde las intensidades se agrupan (de ahí el hecho de llamarlos de *clustering*) en dos grupos distintos, lo cual se suele modelizar como si

se intentasen clasificar esos dos grupos siguiendo dos distribuciones gaussianas, una para cada grupo. Para que este tipo de umbralización dé buenos resultados es importante que el histograma sea suficientemente *bimodal* [15]. Esto implica que el histograma⁵ debe representar una distribución de las intensidades de la imagen de forma que haya un “valle” bien definido entre ambos grupos. Si este valle es inexistente, los resultados no tendrán por qué ser buenos.

Basadas en entropía: aquéllas que analizan la imagen en base a la entropía de los segmentos de fondo y de objeto, la entropía de la imagen binarizada frente a la imagen original o, en general, cualquier derivado que requiera de análisis de entropía para poder decidir sobre los segmentos resultantes.

Basadas en los atributos de los objetos: tratan de buscar similitud entre atributos entre la imagen original y la imagen binarizada resultante tales como el análisis de la forma usando lógica borrosa (*fuzzy*), análisis de la coincidencia entre los bordes, etc.

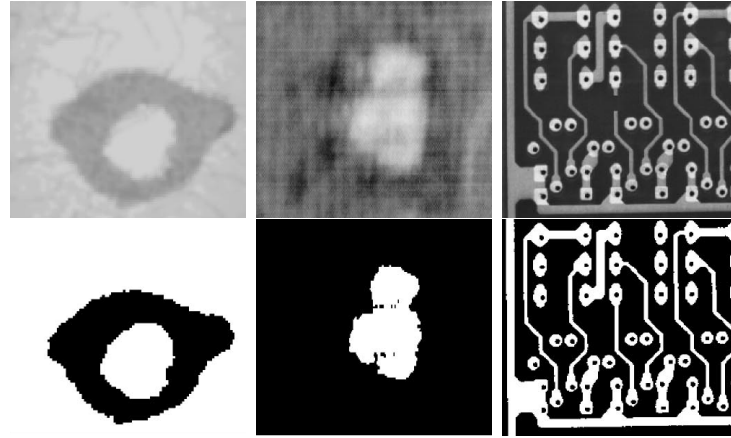
Métodos espaciales: utilizan distribuciones de probabilidad y/o correlación entre píxeles de la imagen para determinar los segmentos.

Métodos locales: este tipo de métodos en lugar de dar un umbral para todos los píxeles, dan un umbral para cada píxel basándose en características locales (aquéllas que pertenecen a los vecinos de los píxeles analizados).

Estas técnicas de umbralización son también populares fuera del ámbito de las imágenes médicas. Pueden ser útiles por ejemplo para automatizar tareas de *testing no destructivo* (NDT) [20]. Puede ser útil en tareas como la detección automática de pequeñas roturas y corrosión del fuselaje de los aviones en imágenes obtenidas mediante corriente Eddy (figura 3.11a). También es común la detección de defectos en plásticos reforzados con fibra de vidrio (GFRP) a través de imágenes de ultrasonidos (figura 3.11b) y en la inspección de circuitos impresos (figura 3.11c). Estas tareas son muy tediosas y requieren de mucha concentración por parte de los operarios que lo hacen manualmente.

Es también muy común el uso de umbralización para la *binarización de documentos*, es decir, la separación del documento en dos segmentos: el fondo, que corresponde al papel; y el objeto que en este caso corresponde a los caracteres que componen el documento [20]. Este tipo de aplicaciones puede servir para la automatización de tareas de lectura de documentos y su conversión a formato electrónico sin necesidad de un operario (figura 3.11d). Además de las imágenes originales, se muestran las imágenes ya umbralizadas manualmente con el umbral ideal para cada caso.

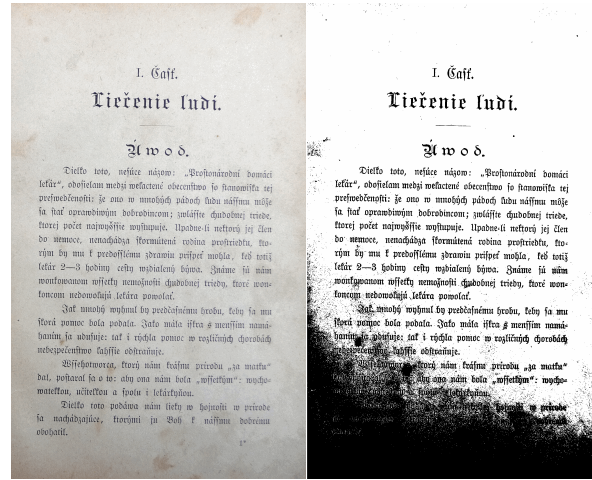
⁵Representación de la distribución de intensidades de la imagen en una serie de rangos establecidos.



(a) Eddy

(b) GFRP

(c) PCB



(d) Binarización de un documento y el resultado tras umbralizar

Figura 3.11: Ejemplos de imágenes umbralizadas en NDT [20] y binarización de documentos

Como estas técnicas se crearon con una visión general, se pueden aplicar a casi cualquier tipo de imagen y pueden dar buenos resultados en general. Todo esto se puede hacer siempre y cuando la imagen cumpla con los requisitos necesarios para poder hacer umbralización, es decir, que la imagen tenga un solo canal de intensidad. Si la imagen tuviese varios canales, habría que convertirlas primero a imágenes de un solo canal.

Estas técnicas se han aplicado ya en nuestro campo, las imágenes PET, aunque el estudio se hizo en seis *phantoms* esféricos de distintos tamaños y distintos niveles de

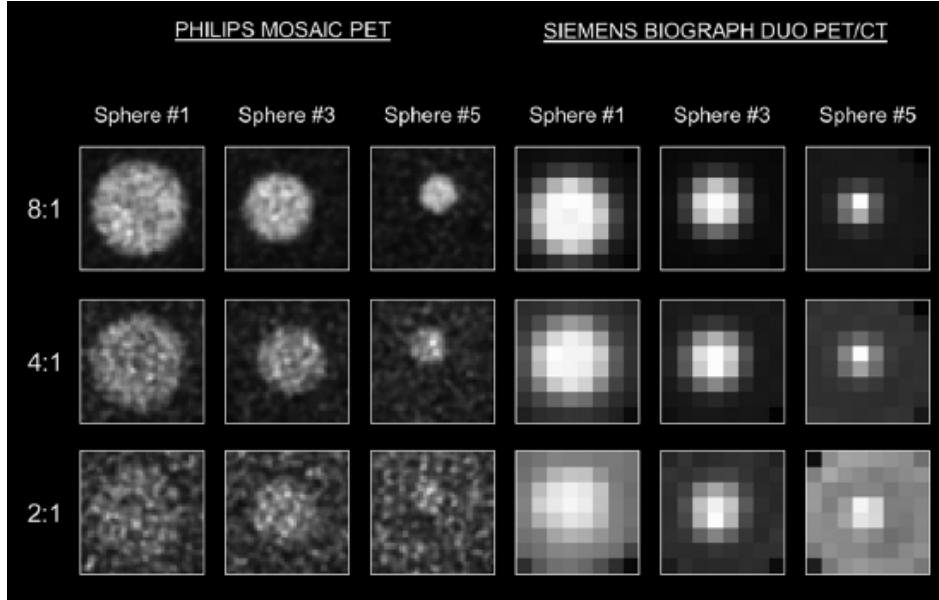


Figura 3.12: Esferas 1, 3 y 5 en ambos escáneres y los niveles de contraste [17]

contraste en cada prueba. Las imágenes PET de estos *phantoms* se obtuvieron usando dos tipos de escáner distintos que tienen distintas características y producen distintas cantidades de ruido al extraer las imágenes [17].

Para la realización de estas pruebas se siguió el protocolo estándar en cada uno de los escáneres, lo que incluye un proceso de corrección de la atenuación para la obtención de las imágenes finales. Para simular el contraste que se da en imágenes PET ordinarias se inyectó marcador de fondo en el material de prueba de ambas pruebas (las de los dos escáneres) en distintas cantidades siguiendo tres niveles de contraste (también conocido como *source-to-background*) diferentes (2:1, 4:1 y 8:1). Además, la profundidad de las imágenes obtenidas se pasó de 16 bits a 8 bits [17].

En resumen, en este experimento se generaron un total de 36 pruebas (6 esferas, 3 niveles de contraste y 2 escáneres) las cuáles se pueden observar en la figura 3.12.

En este experimento se probaron concretamente 21 técnicas que aparecen en la evaluación (*survey*) de técnicas de umbralización del artículo [20]. Primero probando estas técnicas para descartar aquéllas que daban resultados inaceptables y así continuar el experimento con las técnicas no descartadas. De todas esas técnicas, sólo se seleccionaron 12, las cuáles fueron puestas a prueba y cuyos resultados fueron analizados y contrastados posteriormente con la técnica de umbralización fija del 42 % [17].

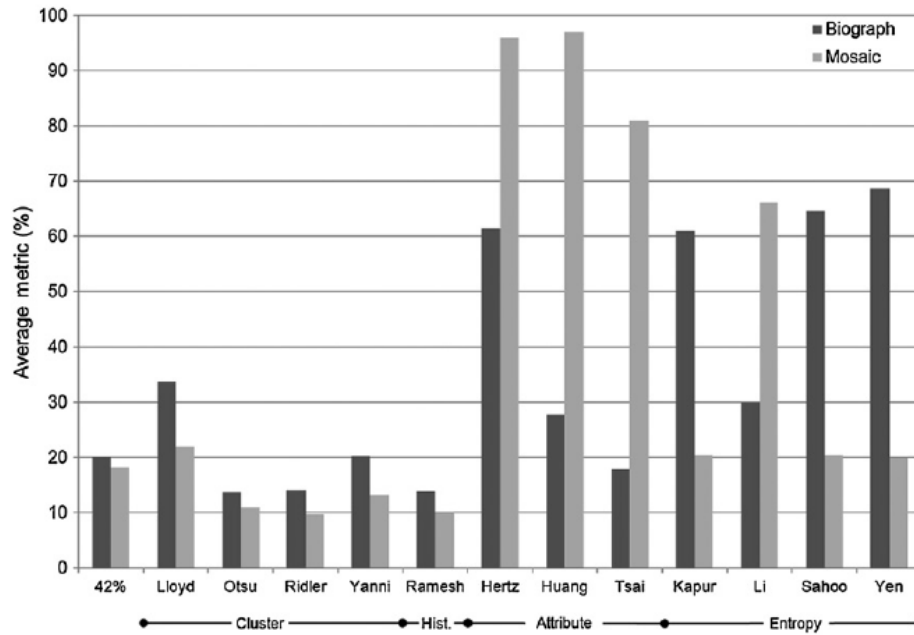


Figura 3.13: Medida de rendimiento de las 12 técnicas probadas [17]

Se utilizaron tres indicadores de rendimiento para comparar los resultados obtenidos con los volúmenes conocidos de las esferas. De estos indicadores se calculó la media para que sirviese de medida de rendimiento definitiva. Según este artículo, las técnicas de *clustering* (ver la clasificación 3.4.1 para recordar cuáles son las técnicas de *clustering*) fueron las que dieron los mejores resultados y, en concreto, las técnicas de umbralización de Ridler, Ramesh, Otsu y Yanni dieron mejores resultados que la técnica de 42 % fijo (ver figura 3.13). De hecho, las técnicas de Ridler y de Ramesh fueron las que mejores resultados dieron [17]. Esta medida de rendimiento está basada en una medida del error, por tanto, cuanto más bajos sean los resultados de los indicadores, mejor es el resultado (de nuevo, ver figura 3.13).

Este artículo es el único artículo encontrado que parece hacer una comparativa completa de técnicas clásicas, aunque sólo se centre en la umbralización.

Umbralización dinámica de PET

Ahora nos centraremos en la umbralización dinámica creada exclusivamente con la intención de ser útil en PET. Estas son técnicas que en lugar de ser de carácter general, se basan en estudios estadísticos y que hacen uso de parámetros tales como el SUV

máximo, el valor de SUV medio de los volúmenes de cáncer obtenidos en experimentos previos o la relación entre la intensidad de los valores de fondo y la intensidad de los valores del tumor o tumores cancerígenos.

A menudo estas técnicas se basan en estudios estadísticos hechos mediante *phantoms*, que, como hemos visto, son volúmenes experimentales expresamente diseñados para hacer estudios de este estilo. En pocas ocasiones se han hecho estudios utilizando casos reales y además, los estudios que se han hecho sobre casos reales utilizan volúmenes hallados manualmente por especialistas para contrastar con los volúmenes obtenidos usando técnicas de segmentación. Evidentemente, estos volúmenes usados para contrastar no tienen por qué ser buenos ya que, como se ha visto en la sección 3.1.3, son subjetivos. Entonces, no hay garantías de que el estudio sea del todo acertado si la validación se basa en resultados manuales [24]. Estos estudios mediante *phantoms* pueden tener ruido y simular la heterogeneidad de los casos reales, aunque evidentemente esto no es suficiente ya que los casos reales tienen muchísima complejidad [24].

En resumen, se han propuesto muchas técnicas de umbralización que utilizan parámetros obtenidos experimentalmente pero cuyos umbrales varían en base de ciertas características generales de la imagen. Estos parámetros son fijos y se han obtenido para maximizar la precisión de estas técnicas en la delineación de PET. Entonces, ninguna de las siguientes técnicas está pensada para ser aplicadas en un ámbito general y lo más probable es que los resultados no sean ni siquiera aceptables si se usan en otras áreas.

Ahora se revisarán brevemente algunas de las técnicas de umbralización dinámicas citadas en el documento [24]:

Relación logarítmica con el volumen del tumor: según Biehl et al. (2006) [3] existe una relación logarítmica entre el volumen del tumor y el umbral fijo relativo, obtenido mediante el estudio de volúmenes delineados manualmente en CT en pacientes con NSCLC. Ver la fórmula 3.4.

$$\text{umbral}(\%) = 59,1 - 18,5 \times \log_{10}(\text{volumen del tumor}) \quad (3.4)$$

Aunque en el artículo [24] no queda especificado en qué unidad se debe introducir el volumen del tumor, los volúmenes con los que suelen trabajar están en mililitros y lo más probable es que se trate de esa unidad.

Basado en la intensidad media y la de fondo: propuesto por Nestle et al. (2005) [13], esta fórmula utiliza la intensidad media de todos los píxeles rodeados por el con-

torno generado sobre el 70 % de la intensidad máxima del tumor y la intensidad media de un órgano vecino suficientemente alejado del tumor como para no verse afectado por su intensidad. La primera intensidad media, se nota en la fórmula como I_{medio} y la intensidad considerada de fondo se nota como I_{fondo} . La fórmula quedaría entonces como sigue en 3.5.

$$\text{umbral}(\%) = \beta \cdot I_{medio} + I_{fondo} \quad (3.5)$$

donde se hayó que el β ideal era de 0.15.

Posteriormente se revisó la fórmula anterior y se dio una nueva versión optimizada para un nuevo conjunto de prueba que incluía *phantoms* esféricos de distintos diámetros. La fórmula resultante fue la 3.6.

$$\text{umbral}(\%) = (a \times SUV_{medio(70\%)} + b \times BKG) / SUV_{\text{máx}} \quad (3.6)$$

En este caso los valores a y b dependen del tamaño de la esfera y del escáner. Concretamente, indicaron que para esferas superiores a 3 cm de diámetro los valores eran $a = b = 0,5$ y que para esferas inferiores a 3 cm de diámetro eran $a = 0,67$ y $b = 0,6$.

Además de estos métodos, se han propuesto otros métodos iterativos que relacionan el volumen de la lesión con otros parámetros como el *source-to-background* y el volumen directamente con el umbral ideal basándose en el método de Monte Carlo [24]. Pero éstos y otros métodos se han omitido para no alargar innecesariamente este apartado.

Como se puede observar, todos estos métodos tienen demasiados parámetros que se deben optimizar y que dependen del estudio que se esté haciendo. De hecho, están muy enfocadas al tamaño de los *phantoms* en muchos casos y esto aplicado en un caso real implicaría conocer el tamaño del tumor cancerígeno cuando esto es precisamente lo que se quiere hallar. Por tanto requieren de demasiados conocimientos *a priori* sobre la lesión y, en muchos casos, carecen de una explicación profunda y razonada de cómo y por qué se deben usar exactamente las fórmulas que proponen.

3.4.2. Técnicas de multiumbralización

La multiumbralización fue nombrada en la sección 3.2 donde se describieron brevemente todas las técnicas de segmentación, entre ellas las técnicas de umbralización. La multiumbralización difiere de la umbralización simple en que se utiliza más de un umbral para segmentar y, por tanto, se obtienen más segmentos de objeto que en la simple.

De nuevo, asumiendo un solo canal de intensidad en la imagen de entrada, podemos describir la multiumbralización usando la siguiente fórmula 3.7.

$$g(\bar{x}) = \begin{cases} 0 & \text{si } f(\bar{x}) < u_1 \\ 1 & \text{si } u_1 \leq f(\bar{x}) < u_2 \\ \dots & \\ n-1 & \text{si } u_{n-1} \leq f(\bar{x}) < u_n \\ n & \text{si } u_n \leq f(\bar{x}) \end{cases} \quad (3.7)$$

Ahora en lugar de tener sólo valores 0 y 1, tenemos que usar más valores para etiquetar los píxeles o vóxeles que corresponda. En esta fórmula tenemos n umbrales. Y de nuevo, al igual que en la umbralización simple, podemos encontrarnos con un caso en el que queramos hacer umbralización local y multiumbralización al mismo tiempo. En este caso, necesitaremos múltiples umbrales descritos para cada píxel o vóxel independientemente. De nuevo, habría que sustituir los u_1, u_2, \dots, u_n por $u_{h_1(\bar{x})}, u_{h_2(\bar{x})}, \dots, u_{h_n(\bar{x})}$ donde cada $h_k(\bar{x})$ sería una función que nos devuelve el índice del umbral que haya que aplicar al píxel o vóxel en la lista de umbrales para \bar{x} . Además, en este caso, habría que añadir la condición de que la lista de umbrales sea estrictamente creciente para evitar repeticiones de umbrales y para que la fórmula anterior tenga sentido.

La multiumbralización es entonces útil para segmentar imágenes donde hay más de un grupo de intensidades de los que estamos interesados en separar. Este tipo de segmentación puede ser bastante útil de nuevo en imágenes NDT como las presentadas en figura 3.11 [21] y parece ser útil para la detección de venas en imágenes de retina [10].

Por otro lado, no ha sido posible encontrar artículos en los que se haya utilizado este tipo de técnica anteriormente en el ámbito de las imágenes PET y CT. Aunque esto no significa necesariamente que sean técnicas que no se puedan adaptar para su uso al menos en PET.

Al no haber conseguido encontrar ninguna referencia sobre multiumbralización, tan sólo podremos introducir ciertas técnicas de multiumbralización clásicas de uso general tal y como se hizo en la sección 3.4.1 con las técnicas clásicas. Para ello nos basaremos en el artículo [19].

En este artículo se señala que los métodos clásicos que vamos a listar a continuación se pueden extender al caso de multiumbralización.

Otsu: técnica de umbralización basada en *clustering* que utiliza la media y las desviaciones típicas de una distribución gaussiana conceptual por cada segmento para

hacer la clasificación de manera similar a otras técnicas basadas en *clustering*. Esta técnica se explicará con más detalle en la sección 4.1.1 ya que fue elegida para hacer nuestras pruebas.

Pun: la cual está basada en entropía y utiliza unas funciones de maximización para obtener el umbral ideal.

Kapur: esta técnica utiliza de nuevo el concepto de entropía, pero esta vez utilizando probabilidades de distribución de las distintas intensidades de la imagen sobre otra función de maximización distinta.

Preservación del momento: en esta técnica los valores de los umbrales se calculan de forma determinista de forma que los momentos de la imagen original se preserven en la imagen umbralizada.

Mínimo error: en esta técnica el histograma se ve como una estimación de la función de densidad de probabilidades sobre la población de píxeles que comprenden el segmento de objeto y el segmento de fondo.

Además de esos métodos, en el artículo se añaden otros métodos que parecen no estar clasificados con los demás métodos probablemente por no ser tan conocidos y ser creados con el objetivo de ser utilizados sólo para multiumbralización.

Método de amplitud de la segmentación: en este método se examina la curvatura de la función de distribución acumulativa para establecer los umbrales.

Método de Wang y Haralick: esta técnica está basada en la distinción de los píxeles que forman bordes en la imagen en base a la intensidad de los píxeles vecinos para establecer un primer umbral, después se aplica sucesivas veces sobre los píxeles clasificados en un segmento y en los píxeles del otro segmento individualmente si es que se quieren obtener más umbrales sobre esos píxeles.

Método de contraste uniforme: esta técnica se basa en la selección recursiva tratando de buscar el umbral que encuentre la mayor cantidad posible de bordes de mayor contraste y la menor cantidad posible de bordes de menor contraste en cada iteración.

En este artículo se hicieron varias pruebas de rendimiento, sin embargo, las imágenes que utilizaron son imágenes comunes del día a día y no serán presentadas aquí puesto que la calidad de las mismas no es muy conveniente dado que el documento del que disponemos parece haber sido escaneado.

Capítulo 4

Metodología

4.1. Técnicas escogidas

Las técnicas que se han escogido para resolver el problema han sido ya introducidas en el capítulo 3 del estado del arte. Las técnicas se han escogido en base a los siguientes criterios por orden de preferencia:

- Debe haber sido preferiblemente utilizada en el campo de delineación de tumores cancerígenos en imágenes PET y haber dado buenos resultados. En su defecto, debe poder ser fácilmente adaptable a nuestro campo de interés.
- Debe requerir la menor cantidad posible de parámetros, es decir, ser lo más parecido a una técnica completamente automática (ver 3.2).
- Debe ser suficientemente sencilla en dos sentidos: debe ser suficientemente rápida en la obtención de los resultados y, debe ser sencilla conceptualmente puesto que eso implicará que será más sencillo razonar sobre lo que está sucediendo si algún resultado difiere de lo esperado.
- Este criterio va en relación con el punto anterior, la técnica debe utilizar conceptos conocidos o, si no se conocen, deben estar bien detallados para evitar ambigüedades. Si la técnica contiene partes que son complicadas de entender, la técnica es descartada al no haber garantías de que la implementación se vaya a corresponder con lo que supuestamente se esté describiendo en el artículo.
- Preferiblemente, y en relación con el primer punto descrito, si la técnica fue probada en el área de delineación en PET, las pruebas deberían ser variadas y deberían

utilizar indicadores suficientemente variados. A continuación se describen en más detalle estas preferencias sobre la “variabilidad” de las pruebas.

- Las pruebas se deben haber realizado preferiblemente con *phantoms* de distintos tamaños, al ser posible, *phantoms* complejos que simulen casos reales, no necesariamente esféricos.
- Las pruebas deben incluir distintos contrastes entre el marcador del tumor y el marcador de fondo para simular la situación real.
- Las pruebas deberían haber sido realizadas también en pacientes reales contrastando los volúmenes obtenidos con volúmenes obtenidos manualmente por los especialistas.

Por esta razón, las técnicas de umbralización son las que se han escogido ya que tienen las siguientes ventajas y desventajas:

■ Ventajas:

- Han sido las más utilizadas en el área de PET [24]. Esto implica que hay numerosos estudios y, por tanto, información completa sobre éstas.
- Son sencillas en ambos sentidos, en tiempo de ejecución y conceptualmente [17].
- La mayoría son reproducibles dado que no hay variaciones respecto de una prueba a otra. Se obtienen siempre los mismos resultados.

■ Desventajas:

- Sensibles a los artefactos producidos por el movimiento [24].
- Problemáticas cuando los tumores son muy heterogéneos [24, 12].

Sin embargo, las desventajas que tienen estas técnicas las podemos encontrar también en otros tipos de técnicas locales basadas por ejemplo en el gradiente, las cuales son sensibles al ruido [24]. No obstante, la heterogeneidad de las lesiones y los artefactos producidos por el movimiento no difieren demasiado visualmente de los efectos producidos por el ruido. Aunque no hay que confundir la heterogeneidad natural del tumor con la heterogeneidad debida al ruido, en definitiva, siempre nos encontramos con ruido en torno a los tumores y también en sus alrededores.

Además, no hay todavía ninguna preferencia en la literatura sobre qué técnicas son definitivamente mejores para las tareas de delineación en hospitales, esto todavía está siendo activamente investigado [24].

Aunque lo que nos interesa es obtener el volumen más ajustado posible, en caso de error, es preferible que el volumen delineado resultante sea ligeramente superior en lugar de ser ligeramente inferior. Esto es así puesto que, hay que recordar que el objetivo es refinar posteriormente la delineación en CT lo cual se hará probablemente modificando el volumen obtenido y adaptándolo a la imagen en CT siendo más fácil reducir dicho volumen que ampliarlo.

Dada la naturaleza de las imágenes PET, las técnicas de multiumbralización podrían ser bastante útiles puesto que idealmente, podrían ayudar a separar automáticamente zonas del fondo de la imagen, del cuerpo, de los tumores y de la vejiga del paciente. Todas estas zonas que han sido nombradas tienen intensidades muy distintas y no tendría por qué ser improbable que la separación de las mismas diese buenos resultados.

4.1.1. Umbralización de Otsu

La técnica que se va a revisar ahora se conoce como umbralización de Otsu y se puede ver en mayor detalle en el artículo original de Otsu ([15]). La técnica de umbralización de Otsu se basa en *clustering* y trabaja directa y explícitamente con distribuciones gaussianas y sus propiedades. Hay que recordar además que por ser una técnica de umbralización basada en *clustering* el rendimiento es mayor cuanto más bimodal sea el histograma (ver 3.4.1). Aunque como se indicó en la sección 3.4.2, esta técnica se puede extender al caso de multiumbralización, con lo cual, el histograma en estos casos debería presentar varios “valles” claramente definidos y preferiblemente no muy anchos para ser más sencilla la detección del umbral en el punto más bajo del valle (equivalente a ser bimodal pero en lugar de “bi-” podríamos llamarlo multi-modal). En la figura 4.1a se muestra un ejemplo de histograma con un valle claramente definido en el cual podemos apreciar sin problemas el umbral ideal y, en la figura 4.1b se muestra otro histograma pero que esta vez, a pesar de ser bimodal, el valle es demasiado amplio y es más complicado determinar cuál sería el umbral ideal.

Para aplicar esta técnica sólo se requiere de una imagen con un canal de intensidad (equivalente a lo que Otsu llama en su artículo *gray-level*) y un histograma que se corresponda con la distribución de los valores de intensidad de la imagen.

Usando la notación de Otsu en dicho artículo, podemos decir que la imagen podrá tener L valores distintos que son representados como $[1, 2, \dots, L]$. Además, llamamos n_i al número de píxeles o vóxeles de la imagen con el valor i . Si esto es así, número total de píxeles es $N = n_1 + n_2 + \dots + n_L$. Así podemos definir cada uno de los *bins*¹ del

¹Cada uno de los rangos de valores sobre los que se hacen los conteos en el histograma.

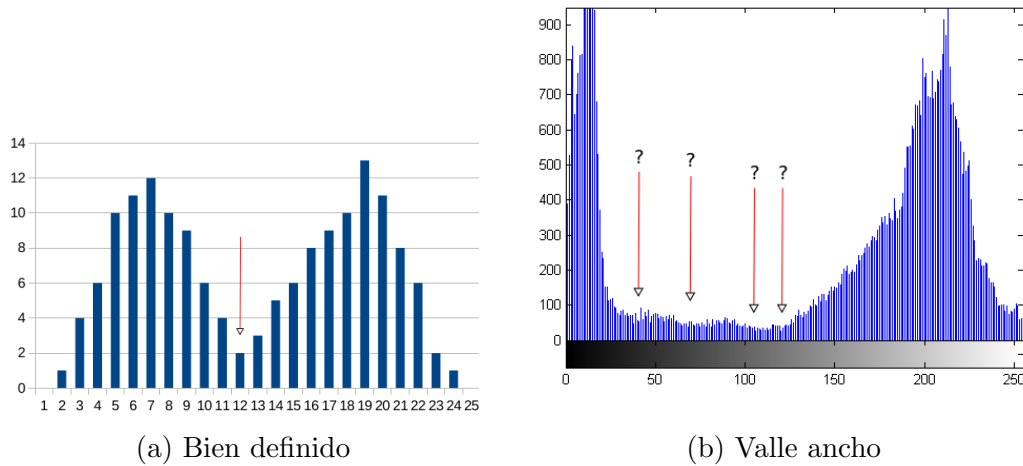


Figura 4.1: Ejemplos de histogramas bimodales

histograma normalizado (sobre N) como sigue en 4.1. Entonces, podemos ver p_i como una probabilidad.

$$p_i = \frac{n_i}{N}, \quad p_i \geq 0, \quad \sum_{i=1}^L p_i = 1 \quad (4.1)$$

Siendo éstas las condiciones de las que partimos, lo que queremos es clasificar los píxeles o vóxeles en dos clases C_0 y C_1 para el segmento de fondo y el de objeto respectivamente. Para hacer esta clasificación tomaremos el umbral k , entonces nos quedará C_0 como los píxeles o vóxeles con intensidades $[1, \dots, k]$ y C_1 como aquéllos con intensidades $[k + 1, \dots, L]$.

Otsu define así la clasificación de píxeles y posteriormente, describe el proceso de obtención de la fórmula final que es la que nos sirve para hacer la umbralización. En este documento, tan sólo daremos la fórmula final y describiremos cada una de las variables que sean necesarias para poder entenderla. Primero definiremos estas variables empezando por las probabilidades de cada clase 4.2 y 4.3.

$$\omega_0 = Pr(C_0) = \sum_{i=1}^k p_i = \omega(k) \quad (4.2)$$

$$\omega_1 = Pr(C_1) = \sum_{i=k+1}^L p_i = 1 - \omega(k) \quad (4.3)$$

Es decir, que ω_0 y ω_1 son respectivamente las probabilidades de que un píxel quede clasificado en la clase C_0 o C_1 respectivamente.

Y por último, tan sólo necesitamos definir las medias en 4.4 y 4.5 que se van a utilizar en la fórmula simplificada final.

$$\mu_0 = \sum_{i=1}^k iPr(i|C_0) = \sum_{i=1}^k \frac{ip_i}{\omega_0} = \frac{\mu(k)}{\omega(k)} \quad (4.4)$$

$$\mu_1 = \sum_{i=k+1}^L iPr(i|C_1) = \sum_{i=k+1}^L \frac{ip_i}{\omega_1} = \frac{\mu(L) - \mu k}{1 - \omega(k)} \quad (4.5)$$

donde $\mu(L)$ se suele notar como μ_T .

Entonces la fórmula final que hay que maximizar es 4.6.

$$\sigma_B^2(k) = \frac{(\mu(L)\omega(k) - \mu(k))^2}{\omega(k)(1 - \omega(k))} \quad (4.6)$$

El k que maximice el valor de $\sigma_B^2(k)$ será el umbral óptimo que será representado como k^* y por tanto se podría formalizar como 4.7.

$$\sigma_B^2(k^*) = \max_{1 \leq k < L} \sigma_B^2(k) \quad (4.7)$$

La técnica de Otsu ha sido escogida por ser una técnica de umbralización de carácter general y porque al parecer, ha dado muy buenos resultados en diversas áreas de NDT y de binarización de documentos tal y como mostraron Sezgin y Sankur en su artículo [20]. Además, esta técnica ha sido probada usando *phantoms* tal y como se describió en la sección 3.4.1 los cuales cumplen al menos la segunda condición sobre los *phantoms* descrita en la sección 4.1.

Podemos afirmar que esta técnica es “suficientemente” automática dado que no requiere ningún parámetros a parte de la imagen de entrada. Por tanto, es una técnica ideal para el objetivo de automatización que queremos cumplir en este proyecto.

Y, por supuesto, al tratarse de una técnica de umbralización, comparte las ventajas y desventajas descritas al comienzo de esta sección 4.1. Esta técnica es sencilla a nivel de conceptos y de optimización debido a que Otsu trató de utilizar el indicador de umbral ideal más simple que pudo encontrar y que utiliza sólo medias, las cuáles son muy sencillas de hallar.

4.1.2. Umbralización de Ridler-Calvard

La técnica de umbralización de Ridler-Calvard es una técnica de umbralización de tipo *clustering* (ver 3.4.1). Se basa en un proceso iterativo de clasificación de píxeles o vóxeles a modo de prueba y error para poder determinar qué hacer en la siguiente iteración. Aquí se resumirá el algoritmo propuesto en el documento original [18] y se dará también una fórmula que se encuentra en este artículo [22] y que facilita la comprensión e implementación del algoritmo.

En esta técnica, Ridler y Calvard propusieron comenzar el algoritmo utilizando una estimación inicial sobre la segmentación de la imagen. Esta estimación inicial consistía en una imagen de las mismas dimensiones que la imagen de entrada, con los 4 píxeles de las esquinas clasificados como segmento de fondo y todos los demás píxeles como segmento de objeto. Concretamente, la figura 4.2 es la imagen que aparece en el artículo.

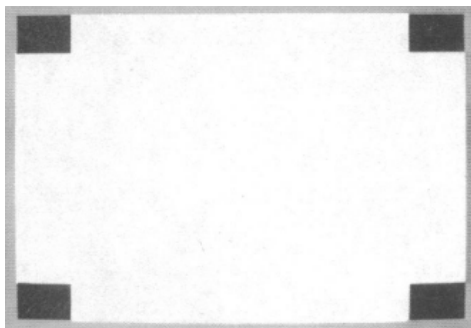


Figura 4.2: Segmentación inicial [18]

El procedimiento a seguir en este algoritmo se describió en el artículo siguiendo un esquema que se ha sido copiado del artículo original figura 4.3.

Según este gráfico, el algoritmo de umbralización de Ridler consiste en introducir la imagen inicial junto con un criterio para la primera iteración, pudiendo usar por ejemplo la propuesta por Ridler y Calvard (4.2), y partiendo de esta imagen inicial, se establece un criterio de clasificación de cada píxel (notado como p en la figura). En las siguientes iteraciones, el criterio que se utiliza es la media de las medias obtenidas en cada segmento. Para hacer estas medias, no hay más que calcular la media, pero sólo de los píxeles pertenecientes al segmento correspondiente.

Este procedimiento se seguiría hasta que no hubiese cambios en el nuevo umbral hallado entre una iteración y la siguiente o hasta un límite en el número de iteraciones.

Para simplificar la comprensión de este algoritmo, se proporciona ahora una fórmula

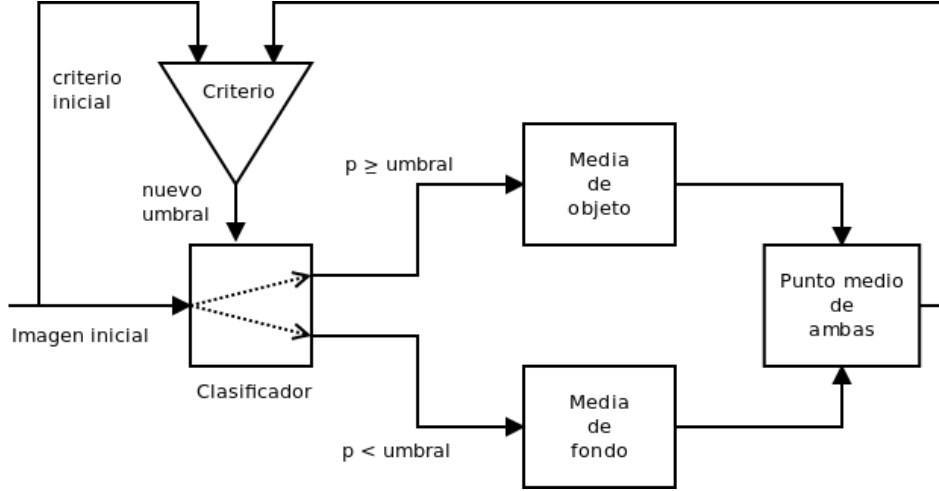


Figura 4.3: Esquema del algoritmo de Ridler [18]

(4.8) obtenida de [22] que resume el procedimiento.

$$T_{k+1} = \frac{\sum_{b=0}^{T_k} bn(b)}{2 \sum_{b=0}^{T_k} n(b)} + \frac{\sum_{b=k+1}^N bn(b)}{2 \sum_{b=k+1}^N n(b)} \quad (4.8)$$

donde T_k es el umbral obtenido en la iteración k -ésima, b es el nivel de “brillo” o lo que conocemos como intensidad de un píxel o vóxel cumpliendo $0 \leq b \leq N$ siendo N la mayor intensidad de la imagen; y $n(b)$ es el número de ocurrencias de la intensidad b , lo cual sería equivalente al número de ocurrencias en el *bin* b del histograma de la imagen.

Esta técnica ha sido escogida por las mismas razones que la técnica de Otsu en todos los aspectos, es decir, es de carácter general y ha dado buenos resultados en diversas áreas [20], ha dado buenos resultados en PET (ver sección 3.4.1), no requiere de parámetros adicionales y, comparte las características de otras técnicas de umbralización. Además, esta técnica en particular es muy sencilla de entender al tratarse de un algoritmo iterativo que utiliza tan sólo medias de las clases de la iteración anterior y, por tanto, es sencilla.

4.1.3. Umbral fijo de Paulino

Esta técnica de umbralización ya ha sido introducida al comienzo de la sección 3.4.1 y dada su simplicidad, no se considera necesaria una explicación en mayor profundidad

ni acerca de la técnica, ni acerca del experimento que llevó a la obtención de este umbral fijo.

Como se dijo también en la sección 3.4.1, parece ser que los umbrales fijos no tienen por qué dar necesariamente resultados buenos y su uso parece que todavía es fruto de controversias. Sin embargo, al tratarse de una técnica tan sencilla y utilizada tan a menudo, parece correcto el empleo de esta técnica en este proyecto para poder hacer una comparativa con las demás que hemos escogido.

4.1.4. Umbralización fija de Erdi

Al igual que sucede en la técnica anterior, ésta es una técnica muy sencilla y, por tanto, no requiere una descripción más a fondo que la que se dio en la sección 3.4.1. Tan sólo hay que especificar el umbral relativo que vamos a utilizar nosotros ya que como dijimos, en principio se podría usar cualquier umbral entre el 36 % y el 44 %. Los valores que se usan más habitualmente en la literatura suelen ser del 40 % (nombrado en [24]) y del 42 %. Como el que se utiliza en el artículo de Prieto et al. (2012) [17] parece ser del 42 %, nosotros haremos lo mismo.

Esta técnica ha sido escogida puesto que asumiendo que usásemos la constante recomendada (42 %) fija, no requiere de más parámetros y ya ha sido empleada en otras ocasiones, esta vez sólo en PET, dando buenos resultados.

De la misma manera que sucede con el umbral fijo de Paulino, el método de Erdi parece ser muy popular y se ha nombrado en diversos artículos sirviéndonos como medida de referencia sobre técnicas comunmente utilizadas para así poder hacer una comparativa. Además, esta técnica ha sido diseñada exclusivamente para PET.

No se han escogido las otras técnicas de umbralización dinámica de PET puesto que o bien requieren de más parámetros tal y como sucede con el volumen del tumor en el método de la relación logarítmica, o bien alguno de los valores a utilizar son demasiado subjetivos como sucede con la fórmula de Nestle et al. (2005) [13] (“órgano vecino *suficientemente* alejado”). Además de que esta técnica ha sido ya probada fuera en pacientes reales tal y como se indica en [24] (pacientes de NSCLC).

4.2. Descripción del material de prueba

Los datos de los que se disponía para realizar la parte experimental, eran las imágenes PET-CT de 6 pacientes diagnosticados de cáncer. Estas imágenes han sido proporcionadas por un hospital público de la Comunidad de Madrid.

Además, se nos ha proporcionado el resultado de la delineación manual hecha por los especialistas de los que dispongan en el hospital. Esto es lo que utilizamos para comprobar la corrección de los resultados que obtenemos con los algoritmos escogidos.

4.2.1. Formato de las imágenes PET

Las imágenes PET están en el formato DICOM (*Digital Imaging and Communications in Medicine*), conocido como el estándar en el ámbito de las imágenes médicas. Este formato contiene los datos de intensidad de cada uno de los *slices* del paciente y además va acompañado de una serie de metadatos conocidos como *cabecera*. Este formato exige que cada *slice* sea guardado junto con los datos de la cabecera en un fichero por separado.

En las cabeceras se guardan múltiples datos relacionados con la identificación del paciente como su nombre y apellidos, dirección, fecha de nacimiento y su sexo. En nuestro caso, los datos relacionados con el paciente han sido eliminados, es decir, las imágenes se han anonimizado por razones de seguridad y además, son datos completamente irrelevantes para lo que nos ocupa en este proyecto.

Por otro lado, se guardan datos relacionados con el hospital como el nombre del mismo y el nombre y apellidos del especialista que le haya hecho las pruebas. Y, unos de los datos más importantes, las cabeceras guardan información sobre el estudio en cuestión como resolución de las imágenes, profundidad del canal de intensidad, posición del paciente, marca del escáner, versión del software del escáner y otras muchas características que son muy importantes para posteriores revisiones.

Vamos a revisar algunas características importantes de nuestras imágenes PET en concreto. Estas imágenes PET tienen una resolución de 168×168 píxeles y el número de *slices* varía de un paciente a otro, siendo 174 slices el que tiene menor número y 198 el que tiene mayor número. Además, también se conocen las dimensiones de cada vóxel, siendo en todos los pacientes de $4,0728 \text{ mm} \times 4,0728 \text{ mm}$ en los ejes X e Y y de 5 mm de grosor de cada *slice*, es decir, la distancia entre planos transversales (eje Z).

Estas imágenes han sido guardadas con una profundidad de 16 bits.

También se incluyen diversos datos en la cabecera de las mismas como el factor de decaída y el peso del paciente, que nos permiten hallar el SUV y así poder transformar toda la imagen de un canal de actividad del marcador a un canal del nivel de SUV representado en forma de números positivos en coma flotante de doble precisión.

Además, hay que añadir que se nos han proporcionado versiones de las imágenes a las que llaman “corregidas” que, al parecer, son las mismas imágenes que las imágenes originales obtenidas por el escáner pero tras haber aplicado un filtro gaussiano o algún otro método de suavizado similar. Desgraciadamente, no sabemos qué método han utilizado exactamente y no ha habido posibilidad de averiguarlo. Daremos por hecho que lo único que se ha hecho a estas imágenes “corregidas” es aplicar métodos de eliminación de ruido.

4.2.2. Características de los pacientes

Los pacientes que se van a estudiar en este proyecto tienen tumores cancerígenos que varían en gran medida tanto de tamaño como de localización y de número de volúmenes disjuntos. Además, también se pueden observar variaciones de intensidad e inhomogeneidad en las lesiones y prácticamente en el resto del cuerpo.

En la tabla 4.1, se presentan los valores del SUV máximo de cada paciente para que sirva de referencia y se puedan apreciar los diferentes rangos de valores que pueden tomar los vóxeles en cada paciente.

Nº de paciente	SUV máximo
0	22.4889
1	34.9941
2	25.6329
3	54.336
4	50.8571
5	44.5358

Cuadro 4.1: Tabla de SUV máximo

Y en la tabla 4.2, se muestran los slices que ocupan los volúmenes definidos por los especialistas para que sirvan de referencia y para saber la situación de los pacientes que disponemos. Tan sólo se indicarán los rangos o intervalos de aquellos *slices* en los que hay al menos un vóxel segmentado como tumor cancerígeno por los expertos.

Los intervalos son presentados además en el mismo orden en que han sido anotados y etiquetados. Cada etiqueta se corresponde con un volumen disjunto (tumor aislado de los demás tumores) al volumen de los demás tumores cancerígenos detectados por el especialista en el paciente. Entonces, en la tabla se muestra también el número de tumores detectados que aparecen en la cabecera de las imágenes DICOM.

Nº de paciente	Nº tumores	Intervalos de slices afectados y etiqueta
0	1	[56, 58] ¹
1	6	[58, 59] ¹ , [57, 59] ² , [71, 74] ³ , [63, 67] ⁴ , [69, 74] ⁵ , [81, 89] ⁶
2	2	[66, 69] ¹ , [76, 87] ²
3	6	[154, 156] ¹ , [85, 86] ² , [84, 87] ³ , [80, 83] ⁵ , [80, 84] ⁶
4	6	[85, 89] ¹ , [82, 87] ² , [71, 73] ³ , [73, 77] ⁴ , [91, 94] ⁵ , [81, 99] ⁶
5	5	[96, 97] ¹ , [93, 97] ² , [77, 80] ³ , [82, 87] ⁴ , [83, 93] ⁵

Cuadro 4.2: Tabla de slices afectados de los pacientes

Como se puede observar en la tabla, en el paciente 3, hemos puesto 6 tumores y sin embargo, se presentan tan sólo 5 intervalos. Esto sucede porque en la cabecera aparecen 6 tumores, pero uno de ellos no tiene ningún vóxel anotado. Entonces, en las anotaciones manuales del paciente 3 distinguimos los volúmenes con etiquetas 1, 2, 3, 5 y 6 (la 4 es la que corresponde al volumen vacío). Con lo cual, simplemente hemos descartado dicho volumen en nuestras pruebas.

Además de este problema, tenemos otros problemas con el paciente 4 ya que hay un solapamiento de volúmenes y la existencia de un par de volúmenes anotados como disjuntos pero que realmente son conexos en las versiones manuales.

Nosotros estamos considerando dos volúmenes como conexos cuando al menos dos vóxeles pertenecientes a cada uno de los volúmenes son vecinos o lo que es lo mismo, se tocan. La conexión en dos dimensiones se puede considerar de dos maneras, en 4-vecinos y en 8-vecinos, sin y con diagonales repectivamente. En tres dimensiones, los tipos conocidos de conexiones son de 6-vecinos (vecinos por las caras del cubo), 18-vecinos (añadiendo los vecinos de las aristas) y 26-vecinos (añadiendo los vecinos de los vértices). Nosotros estamos considerando el caso de 26 vecinos.

Ahora mostramos estos problemas de los que estamos hablando con el paciente 4. En la figura 4.4 podemos observar que los volúmenes etiquetados como 1 y 2 se solapan en el slice 86 (cada uno con distinto color). Y en la figura 4.5 podemos ver que los volúmenes 5 y 6 son conexos en los slices 92 y 93 (de nuevo, cada volumen de distinto color).

Tras esta revisión de los pacientes, vamos a mostrar todos los pacientes en tres

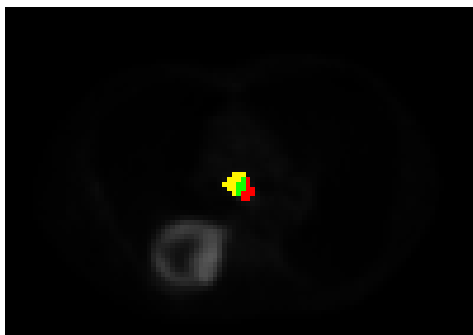


Figura 4.4: Volúmenes 1 (rojo), 2 (amarillo) e intersección (verde)



(a) *Slice 92*

(b) *Slice 92*

Figura 4.5: Volúmenes 5 y 6

dimensiones usando el programa de visualización OsiriX (figura 4.6). Se han añadido notas manualmente para poder indicar qué zonas se corresponden con los volúmenes de tumor cancerígeno obtenidos tras ser anotados manualmente por los especialistas. Las notas consisten en circunferencias y lazos verdes que rodean las zonas en las que están los tumores reales junto con la etiqueta del tumor correspondiente. Así podremos situarnos intuitivamente cuando hablemos de un paciente u otro más adelante. Para apreciar estos volúmenes de la mejor forma posible, se ha colocado cada uno de los pacientes de manera que en la reconstrucción 3D hecha por OsiriX se distingan fácilmente los distintos tonos de las imágenes. En estas imágenes el nivel de SUV es proporcional al mapa de colores que va desde negro para niveles de SUV inferiores pasando por tonos de rojo, naranja, amarillo para niveles de SUV intermedios y finalmente blanco para niveles de SUV superiores.

En el paciente 5, no hemos sido capaces de encontrar en la reconstrucción 3D el tumor etiquetado como 1. Por tanto, este tumor se ha anotado como “1?” en la zona donde se sitúa, aunque no se sepa exactamente a qué vóxeles se refiere. Y en el paciente 4, se han utilizado lazos en lugar de circunferencias para indicar los volúmenes, ya que

están demasiado juntos como para que se vean correctamente.

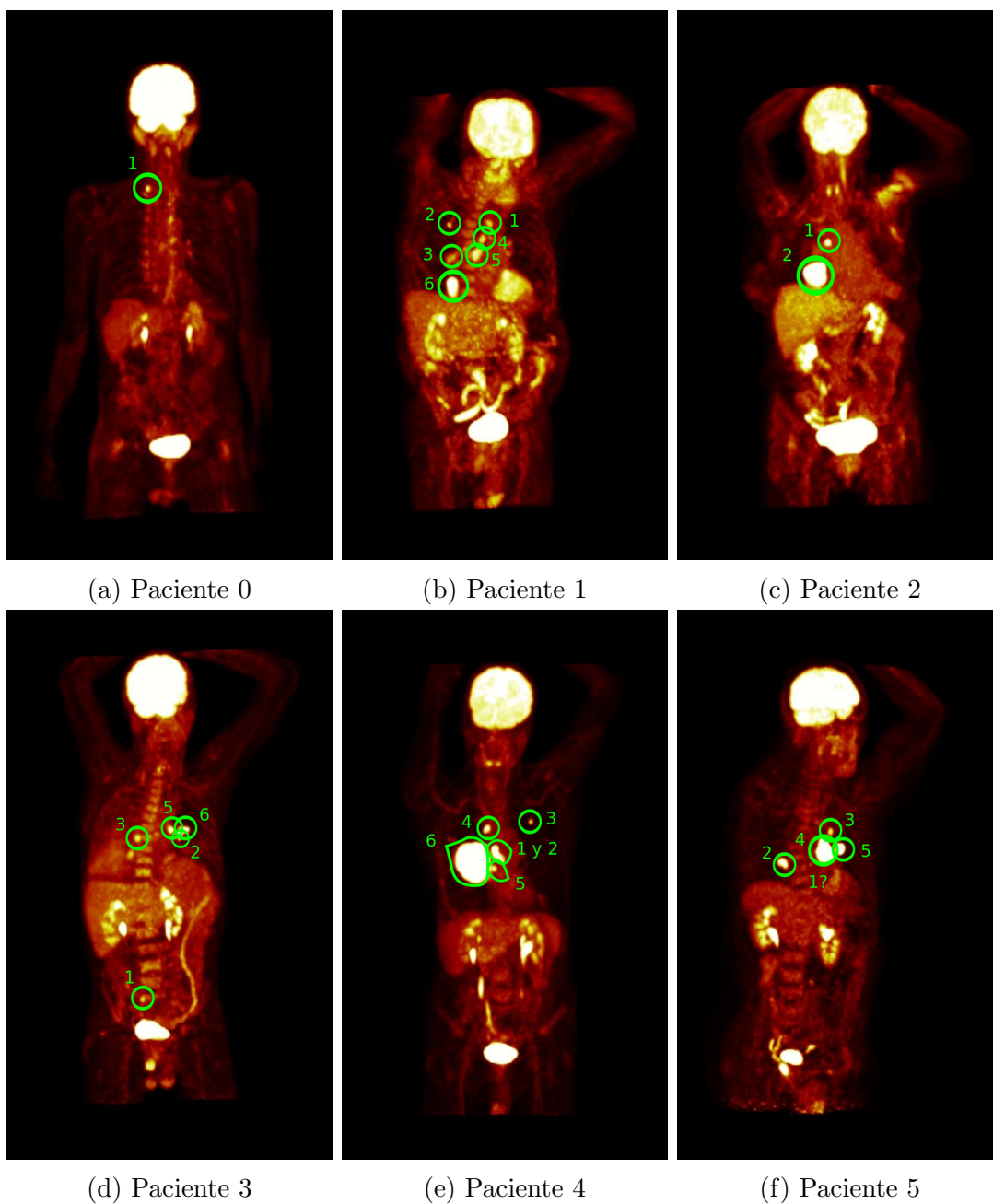


Figura 4.6: Reconstrucción 3D de las imágenes PET

4.3. Métodos de evaluación

Primero vamos a describir brevemente el procedimiento de realización de las pruebas brevemente para poder llevar a cabo la evaluación correctamente. Se mostrará el resumen del procedimiento por orden de ejecución:

Carga de los *slices* del paciente: el paciente de la base de datos se carga y convierte a unidades del SUV.

Proceso de segmentación: el paciente está listo para ser segmentado y se sigue el proceso de segmentación ya descrito en 3.1.2.

Separación de volúmenes no conexos: se separan los volúmenes que sean no conexos usando la relación de conexión de 26-vecinos descrita en la sección 4.2.2, a cada volumen separado se le asigna una etiqueta distinta.

Elección de las ROIs: se eligen manualmente los volúmenes que se corresponden con un tumor ya que conocemos dónde están situados. Recordar que la intención de este trabajo es la delineación de las ROIs, pero no la discriminación de las regiones irrelevantes. Idealmente, este trabajo debería ser realizado por otro módulo en el proyecto presentado en la sección 2.1.

Evaluación final: se hace una evaluación de cada uno de estos volúmenes por separado siguiendo los métodos de evaluación que vamos a describir a continuación.

Los métodos de evaluación que se han elegido para evaluar el rendimiento de las técnicas escogidas son los mismos que se proponen en el artículo [17]. Estos métodos se describirán a continuación y se dará una motivación de su uso. Pero primero se dará una guía sobre la nomenclatura que se va a utilizar, siendo la misma que se usa en dicho artículo, ya que es común a todos los métodos. Éstas son las medidas que se usarán en el capítulo 5.

F_0 = conjunto de vóxeles pertenecientes al segmento de objeto real (manualmente delineado por los especialistas).

B_0 = conjunto de vóxeles pertenecientes al segmento de fondo real (manualmente delineado por los especialistas).

F_t = conjunto de vóxeles pertenecientes al segmento de objeto a evaluar (segmento de objeto hallado mediante la técnica).

B_t = conjunto de vóxeles pertenecientes al segmento de fondo a evaluar (segmento de fondo hallado mediante la técnica).

Y ahora vamos a describir los métodos de evaluación con la anterior nomenclatura:

Medida de error sobre el volumen (VE) - ésta es una medida del error que evalúa la relación entre el volumen hallado ($|F_t|$) y el volumen real cuantitativamente ($|F_0|$). Entendemos que el operador $|A|$ sobre el conjunto A es el cardinal de este conjunto, es decir, el número de elementos que tiene. En nuestro caso, el cardinal se corresponde directamente con el volumen ya que es la cantidad de vóxeles que hay en el conjunto. Esta medida se puede representar como muestra la fórmula 4.9.

$$VE(\%) = 100 \cdot \left| \frac{|F_0| - |F_t|}{|F_0|} \right| \quad (4.9)$$

Medida de error sobre la clasificación (CE) - ésta es una medida del error que evalúa la relación entre la cantidad de vóxeles clasificados erróneamente y la cantidad de vóxeles que debería haber. En esta medida, se tienen en cuenta dos casos distintos, la cantidad de vóxeles clasificados como objeto en las ROIs reales que quedan clasificados como vóxeles de fondo en nuestra segmentación y la cantidad de vóxeles clasificados como objeto en nuestra segmentación que quedan clasificados como vóxeles de fondo en las ROIs reales. La ventaja de esta medida respecto de la anterior es que además de incluir el volumen, incluye también el posicionamiento correcto o incorrecto de los vóxeles ya que podemos obtener un volumen idéntico en cantidad pero estar desplazado del volumen real. Esta medida se puede representar como muestra la fórmula 4.10.

$$CE(\%) = 100 \cdot \frac{|F_0 \cap B_t| + |B_0 \cap F_t|}{|F_0|} \quad (4.10)$$

Dice similarity index (DSI) - en este caso, no se trata de una medida de error, sino que se trata de una medida de similitud (también en porcentaje). En este caso, lo que estamos haciendo es medir la similitud entre los dos volúmenes de objeto en cuanto al número de coincidencias que tienen o lo que es lo mismo, el cardinal del conjunto de intersección. Esta medida ofrece información desde otro punto de vista y además, incluye la característica del posicionamiento puesto que se mide la cantidad de vóxeles que coinciden. Sin embargo, esta medida no sirve para comprobar cuántos vóxeles se salen del volumen ideal, por tanto, podemos obtener un resultado del 100 % de similitud y tener un volumen mucho más grande que otro. Esto puede suceder si uno de los volúmenes está contenido en el otro

volumen. Esta medida se puede representar como muestra la fórmula 4.11.

$$\text{DSI}(\%) = 100 \cdot \frac{2 \cdot |F_0 \cap F_t|}{|F_0| + |F_t|} \quad (4.11)$$

Capítulo 5

Resultados

En este capítulo se muestran los resultados obtenidos usando las técnicas escogidas (ver sección 4.1) siguiendo el proceso descrito en la sección 4.3 sin utilizar ningún tipo de preprocesado salvo aquél que se esté usando en las imágenes “corregidas” al ser sobre las que hemos trabajado. Además, se ha utilizado el SUV para hacer las pruebas (revisar sección 3.1.3 para recordar qué es el SUV).

Primero se mostrarán los umbrales obtenidos con cada técnica escogida en la tabla 5.1 salvo en el caso de la técnica de Paulino dado que se trata de un umbral constante sobre el SUV de 2,5.

Pacientes	Otsu	Ridler	42 %
Paciente 0	2,6803	2,6858	9,4441
Paciente 1	13,12	13,1446	14,6975
Paciente 2	6,211	1,7374	10,7658
Paciente 3	15,9929	1,0064	22,82
Paciente 4	3,0559	3,0808	21,36
Paciente 5	2,8284	2,8365	18,7051

Cuadro 5.1: Umbrales obtenidos con cada técnica sobre el SUV

En las próximas tablas (5.2, 5.3, 5.4, 5.5, 5.6, 5.7, 5.8 y 5.9), se muestran los resultados usando el siguiente formato: **Pi-Tj** que quiere decir **P**aciente **i**-ésimo y **T**umor o volumen **j**-ésimo. Así nos referiremos a los resultados obtenidos tras aislar cada uno de los volúmenes disjuntos manualmente en cada uno de los pacientes tal y como se describió en el procedimiento de las pruebas. Además, para referirse a los métodos de

evaluación, se utilizarán las siglas mostradas en la sección 4.3.

Para los casos especiales de ROIs que se han visto en la sección 4.2.2 se ha seguido además del procedimiento general, otro procedimiento distinto. En estos casos especiales, en lugar de evaluar cada volumen obtenido con cada volumen de las ROIs reales (las obtenidas por los especialistas), se han juntado las ROIs reales que cumplían las condiciones de ser conexas o de solaparse. Es decir, se han juntado con una sola etiqueta las ROIs etiquetadas como volúmenes 1 y 2 del paciente 4 y 5 y 6 del paciente 4. Estas ROIs se han notado P4-T1+T2 y P4-T5+T6 respectivamente. Los resultados de los casos especiales, se muestran en las tablas 5.3, 5.5, 5.7 y 5.9.

Esto se ha hecho de ambas formas ya que estaríamos siendo parciales si tratásemos a estas dos ROIs de distinto modo que las demás. No obstante, también sería injusto tratar ROIs conexas como no conexas ya que esto no favorece a los métodos que estamos evaluando y los resultados no serían del todo fieles a la realidad. Por tanto, se presentan ambos resultados.

En las tablas 5.2 y 5.4 se puede observar que los resultados en el paciente 0 han sido perfectos tanto en el algoritmo de Otsu como el de Ridler. Y decimos “perfectos” porque los resultados han sido de 0 % de error en ambos indicadores de error y de 100 % en la similitud entre volúmenes dada su intersección. Éste es el ejemplo de resultado ideal que se está buscando obtener. Todos estos casos ideales han sido resaltados en las tablas.

El peor resultado que podemos obtener se da en aquéllos cuyo error es del 100 % tanto en VE como CE y del 0 % de similitud, significa que directamente no existe dicho volumen, es decir, son casos en los que los volúmenes correspondientes no han sido detectados y por tanto están vacíos.

Los resultados encontrados en la tabla 5.2 del paciente 1 y 3, de la tabla 5.4 o de prácticamente todos los pacientes y todas las combinaciones posibles en las tablas 5.8 y 5.9 cumplen con lo descrito en el párrafo anterior. Lo que significa que no se detectaron los tumores que debían ser detectados, de hecho, no se detectó nada a parte de la vejiga la cual siempre ha aparecido en todos los pacientes debido a su alto nivel de SUV.

En otros casos, aparecen números superiores al 100 %. Estos valores en la fórmula de VE (4.9) significan que la diferencia de volúmenes es superior al 100 % del volumen real. En particular, un valor de 200 % significa que, por ejemplo, un volumen $|A| = 4$ y un volumen $|B| = 12$ nos daría $|4 - 12| = 8$ y en consecuencia $(8/4) \times 100 = 200$. La relación entre esos volúmenes es que el volumen hallado B es 3 veces el volumen real A . Por otro lado, los valores de CE (4.10) superiores a 100 % indican que hay muchos más

Paciente-Tumor	VE(%)	CE(%)	DSI(%)
P0-T1	0	0	100
P1-T1	100	100	0
P1-T2	100	100	0
P1-T3	100	100	0
P1-T4	100	100	0
P1-T5	100	100	0
P1-T6	100	100	0
P2-T1	96	96	7,6923
P2-T2	53,641	53,641	63,3497
P3-T1	100	100	0
P3-T2	100	100	0
P3-T3	100	100	0
P3-T5	100	100	0
P3-T6	100	100	0
P4-T1	147,7612	189,5522	45,4936
P4-T2	71,134	106,1856	60,8365
P4-T3	54,5455	54,5455	62,5
P4-T4	8,5714	25,7143	86,5672
P4-T5	12560	12640	0,9404
P4-T6	2,8843	13,9307	92,9328
P5-T1	50	50	6,6667
P5-T2	8,8235	17,647	90,7692
P5-T3	19,0476	19,0476	89,4737
P5-T4	22,069	22,069	87,5969
P5-T5	5,3846	8,9744	95,3888

Cuadro 5.2: Resultados de evaluación usando la técnica de Otsu

Paciente-Tumor	VE(%)	CE(%)	DSI(%)
P4-T1+T2	1,2195	44,5122	67,2727
P4-T5+T6	3,6236	13,6724	93,0377

Cuadro 5.3: Evaluación de casos especiales usando Otsu

Paciente-Tumor	VE(%)	CE(%)	DSI(%)
P0-T1	0	0	100
P1-T1	100	100	0
P1-T2	100	100	0
P1-T3	100	100	0
P1-T4	100	100	0
P1-T5	100	100	0
P1-T6	100	100	0
P2-T1	96	96	67,5676
P2-T2	1.527,282	1.528,923	11,4839
P3-T1	884.792,3	884.792,3	0,02
P3-T2	2.875.800	2.875.800	0,01
P3-T3	479.216,70	479.216,70	0,04171736
P3-T5	442.346,2	442.346,2	0,04519303
P3-T6	216.949,1	216.949,1	0,09210263
P4-T1	143,2836	185,0746	46,087
P4-T2	68,0412	107,2165	60
P4-T3	54,5455	54,5455	62,5
P4-T4	8,5714	25,7143	86,5672
P4-T5	12.496	12.576	0,9452
P4-T6	3,3753	14,1761	92,7903
P5-T1	50	50	6,6667
P5-T2	8,8235	17,647	90,7692
P5-T3	19,0476	19,0476	89,4737
P5-T4	22,7586	22,7586	87,1595
P5-T5	5,5128	9,1026	95,3197

Cuadro 5.4: Resultados de evaluación usando la técnica de Ridler

Paciente-Tumor	VE(%)	CE(%)	DSI(%)
P4-T1+T2	60,9756	45,122	66,6667
P4-T5+T6	4,1108	13,916	92,896

Cuadro 5.5: Evaluación de casos especiales usando Ridler

Paciente-Tumor	VE(%)	CE(%)	DSI(%)
P0-T1	9,0909	9,0909	95,6522
P1-T1	19,0476	19,0476	91,3044
P1-T2	8,3333	8,3333	95,6522
P1-T3	9,0909	27,2727	85,7143
P1-T4	28,2051	48,718	78,6517
P1-T5	26,3736	26,3736	88,3495
P1-T6	7,9051	9,4862	95,0617
P2-T1	28,00	36,00	79,0698
P2-T2	6,9744	9,2308	95,2179
P3-T1	15,3846	15,3846	91,6667
P3-T2	1550	1550	11,4286
P3-T3	4,1667	29,1667	85,1064
P3-T5	19,2308	19,2308	89,3617
P3-T6	24,5283	32,0755	85,7143
P4-T1	222,3881	234,3284	44,523
P4-T2	122,6804	122,6804	61,9808
P4-T3	0	0	100
P4-T4	14,2857	31,4286	85,3333
P4-T5	14024	14032	1,3498
P4-T6	8,3461	11,2304	94,6097
P5-T1	0	0	100
P5-T2	10,2941	13,2353	93,7063
P5-T3	14,2857	14,2857	93,3333
P5-T4	11,7241	14,4828	92,3077
P5-T5	2,4359	2,9487	98,5434

Cuadro 5.6: Resultados de evaluación usando la técnica de Paulino

Paciente-Tumor	VE(%)	CE(%)	DSI(%)
P4-T1+T2	31,7073	43,9024	71,579
P4-T5+T6	4,1108	13,916	92,896

Cuadro 5.7: Evaluación de casos especiales usando Paulino

Paciente-Tumor	VE(%)	CE(%)	DSI(%)
P0-T1	100	100	0
P1-T1	100	100	0
P1-T2	100	100	0
P1-T3	100	100	0
P1-T4	100	100	0
P1-T5	100	100	0
P1-T6	100	100	0
P2-T1	100	100	0
P2-T2	91,1795	91,1795	16,2111
P3-T1	100	100	0
P3-T2	100	100	0
P3-T3	100	100	0
P3-T5	100	100	0
P3-T6	100	100	0
P4-T1	100	100	0
P4-T2	100	100	0
P4-T3	100	100	0
P4-T4	100	100	0
P4-T5	100	100	0
P4-T6	100	100	0
P5-T1	100	100	0
P5-T2	100	100	0
P5-T3	100	100	0
P5-T4	100	100	0
P5-T5	79,7436	79,7436	33,6887

Cuadro 5.8: Resultados de evaluación usando la técnica de Erdi

Paciente-Tumor	VE(%)	CE(%)	DSI(%)
P4-T1+T2	100	100	0
P4-T5+T6	100	100	0

Cuadro 5.9: Evaluación de casos especiales usando Erdi

vóxeles mal clasificados que la cantidad de vóxeles existentes en el volumen real. En este caso, un 200 % significa que hay una cantidad del 100 % de vóxeles reales que no han sido segmentados como objeto y otro 100 % de vóxeles hallados que no coinciden con los vóxeles reales asumiendo que ambos tuviesen el mismo volumen.

Valores como el caso P4-T4 en la tabla 5.2 indican que el volumen se ha ajustado mucho al caso real (VE bastante bajo) pero que ha habido ciertos vóxeles que no se correspondían, lo que indica cierto desplazamiento (CE del 26 % y DSI del 87 %) con respecto al volumen real.

En muchos casos sucede que los valores de VE y CE coinciden tal y como sucede con P4-T5, P5-T1 o P5-T3 de la tabla 5.2 o todos los valores de VE y CE en los pacientes 2 y 3 de la tabla 5.4. En estos casos lo que está ocurriendo es que el volumen hallado está contenido en el volumen real o viceversa, por tanto, la cantidad de vóxeles clasificados erróneamente coincide con la diferencia de volúmenes y como VE y CE tienen el mismo denominador, se obtiene el mismo resultado. En concreto, los resultados de las evaluaciones obtenidos en el algoritmo de Ridler (ver tabla 5.4) del paciente 3 han sido muy malos dado que el valor del umbral de Ridler en este paciente (ver tabla 5.1) ha sido muy bajo y ha clasificado como tumor casi todo el cuerpo del paciente.

Por último, comentamos los casos especiales de todas las técnicas salvo la de Erdi (ver tablas 5.3, 5.5 y 5.7). En todos estos casos, se puede observar que los resultados han mejorado, especialmente en el caso del tumor 5 del paciente 4 al ser un volumen relativamente pequeño conectado a un volumen relativamente grande (ver figura 4.5). Y evidentemente, al hacer la evaluación, la cantidad de vóxeles mal clasificados era equivalente a prácticamente todos los vóxeles del tumor 6 del mismo paciente lo que eleva el índice de error exageradamente.

De forma general podemos afirmar que el algoritmo de Erdi no ha proporcionado buenos resultados. De hecho, no se ha hallado ningún volumen salvo la vejiga en la mayoría de los casos. Sin embargo, todos estos resultados son lógicos dado que esta técnica se ha utilizado exclusivamente sobre volúmenes cancerígenos que ya se conocían *a priori*. Por tanto, como esta técnica actúa sobre el SUV máximo de toda la imagen, en esos casos, el SUV máximo era el SUV máximo del tumor. Pero en nuestro caso, el SUV máximo es del cuerpo entero lo que implica que la vejiga muy probablemente sea la que tenga ese SUV máximo. Si la vejiga tuviera un SUV máximo y una distribución de los valores de SUV similar al de los volúmenes que nos interesan, esto no habría sucedido. Por tanto, esta técnica no se puede considerar una técnica automática para usar sobre el cuerpo completo del paciente al verse tan afectada por otros SUV superiores a los de los tumores.

De cara a comparar los resultados obtenidos por las técnicas de Otsu y Ridler, podemos concluir que han dado resultados similares en todos los pacientes salvo el 2 y el 3 en base a los umbrales obtenidos y mostrados en la tabla 5.1. Y si analizamos los resultados obtenidos usando el umbral fijo de Paulino, podemos observar que este umbral está en muchas ocasiones por encima de del SUV mínimo en varios de los pacientes. Para ello se muestra la tabla 5.10.

Paciente-Tumor	Volumen	SUV min	SUV max
P0-T1	11	2,7799	5,3753
P1-T1	12	2,4585	4,5358
P1-T2	21	2,5344	4,3713
P1-T3	39	1,8605	3,754
P1-T4	33	1,3767	4,7814
P1-T5	91	2,5418	5,7138
P1-T6	253	1,7291	13,0147
P2-T1	25	1,7382	6,2653
P2-T2	975	1,4778	12,9733
P3-T1	13	1,9169	3,7841
P3-T2	4	2,5836	4,2203
P3-T3	24	1,4278	4,0063
P3-T5	26	1,5173	5,4955
P3-T6	53	2,1972	6,9962
P4-T1	67	1,9432	9,5593
P4-T2	97	2,5	10,399
P4-T3	11	2,5252	4,9263
P4-T4	35	1,9618	7,8815
P4-T5	25	2,1605	6,1944
P4-T6	3259	1,7771	17,9157
P5-T1	4	2,6463	3,1492
P5-T2	68	2,4995	9,4856
P5-T3	21	2,509	5,4774
P5-T4	145	0,5681	18,4344
P5-T5	780	1,8362	44,5358

Cuadro 5.10: Datos de volumen, SUV mínimo y máximo en cada una de las ROIs manuales

A modo ilustrativo, se incluyen algunas imágenes de la reconstrucción 3D de los resultados de la segmentación en los pacientes 0, 2 y 5 (figura 5.2). Se han escogido estas imágenes porque se distinguen claramente los volúmenes que se corresponden con tumores reales. Se han anotado la mayoría de los volúmenes junto con una descrip-

ción breve para que sea más sencilla la asociación de cada uno de ellos con lo que se corresponden en la realidad.

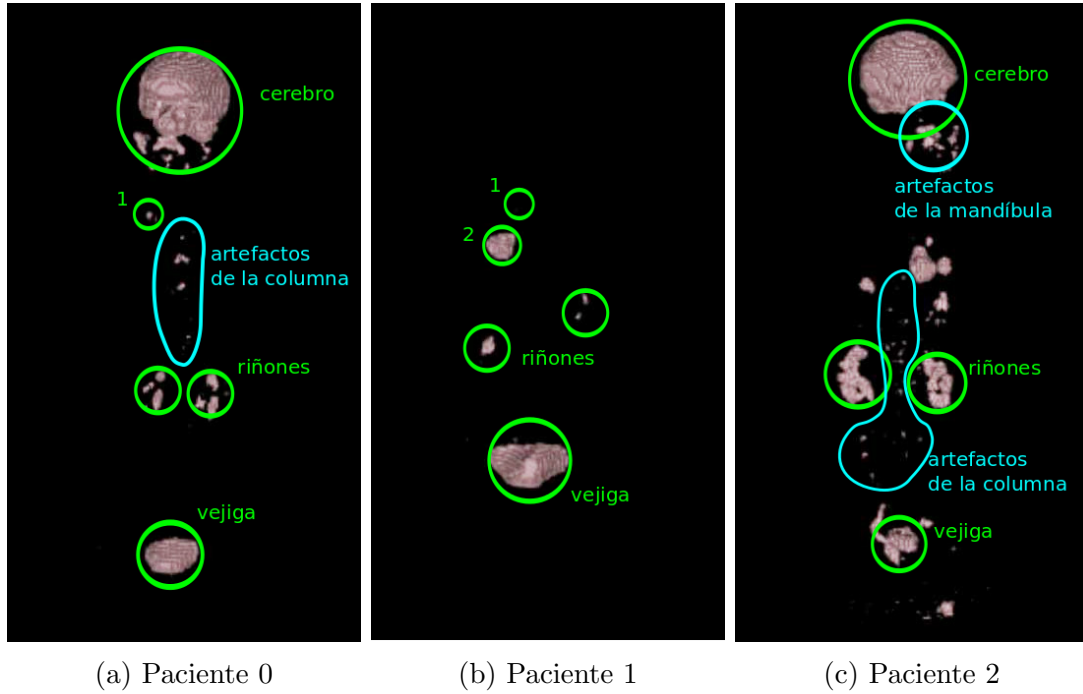
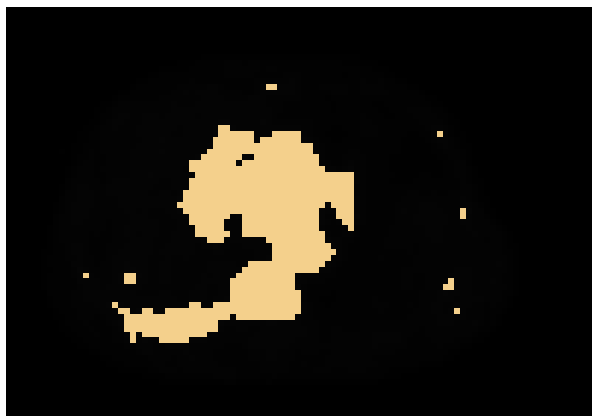


Figura 5.1: Resultados en varios pacientes

En estas imágenes se pueden observar varios volúmenes que se corresponden con ningún tumor cancerígeno real. Entre ellos se encuentran normalmente el cerebro, los riñones y la vejiga. Además de estos volúmenes, aparecen otros derivados de los efectos de ruido en la imagen y que llamamos “artefactos”.

Además de estos resultados buenos, se muestra el *slice* 85 del paciente 3 usando el algoritmo de Ridler. Como se puede ver en la tabla 5.4, es el mejor ejemplo de peor resultado obtenido si no contamos con los resultados usando la técnica de Erdi. Si se observa el umbral obtenido en la tabla 5.1, éste no es muy inferior al SUV mínimo de P3-T3 de la tabla 5.10 y, sin embargo, los resultados no se corresponden con esta relación de umbrales.



(a) Volumen P3-T3 usando Ridler



(b) Volúmenes manuales correspondientes a P3-T3

Figura 5.2: *Slice* 85 del paciente 3

Capítulo 6

Conclusión

En este trabajo, se ha hecho una revisión rigurosa del estado del arte y se han proporcionado resultados de varias técnicas encontradas en la literatura que no habían sido probadas anteriormente en imágenes PET de pacientes reales. Han sido probadas la técnica de umbralización de Otsu y la técnica de umbralización de Ridler. Y, a parte de estas dos técnicas, se han utilizado otras técnicas de umbralización probadas en pacientes reales y más conocidas en la segmentación de imágenes PET para la delineación de tumores cancerígenos que se han usado en este documento para comparar los resultados con las otras dos técnicas.

De todas las técnicas encontradas en la literatura, se comenzaron eligiendo las técnicas de umbralización desde un comienzo dado que existe un alto contraste entre la concentración de marcador que hay en los tumores y la concentración de marcador en el resto del cuerpo y fondo de las imágenes PET. Lo cual favorece a estas técnicas debido a su naturaleza tal y como se ha expuesto en 4.1.

Los resultados obtenidos y expuestos en el capítulo 5, en general, no han sido suficientemente buenos. Aunque tampoco podemos afirmar que estos resultados han sido validados en su totalidad puesto que tan sólo se dispone de 6 pacientes para llevar a cabo la evaluación. Y aunque cada uno de estos pacientes tiene diversos tumores muy variados en tamaño, forma y concentración de marcador, no es suficiente para poder decidir si estas técnicas deberían ser descartadas o no. Idealmente, la mejor opción encontrada en la literatura es la combinación de pruebas usando tanto imágenes de pacientes reales anotadas manualmente como *phantoms* de tamaños, contrastes y concentración de marcador variada. También parece ser una buena opción la utilización de *phantoms* que simulen torsos reales tales como los utilizados en el artículo de Zaidi et al. (2010)

[24]. Parece ser que se ha propuesto el uso de *phantoms* que simulan los movimientos de los pacientes debidos a la respiración durante las sesiones en los escáneres.

Además de estos resultados, en este trabajo se ha aportado una primera aproximación al problema de delineación junto con la seguridad de que muchas técnicas han sido revisadas para hacer la selección de técnicas presentada. Y sobre todo, este trabajo ayudará a futuros estudiantes e investigadores a decidir por qué camino se debería continuar.

También cabe destacar que la mayoría de la literatura encontrada y revisada en este documento trata casos de carcinomas pulmonares no microcíticos (*non-small-cell lung cancer* o NSCLC) y casos de cáncer en el cuello y la cabeza en las que se utilizan imágenes de zonas reducidas. Para el resto de casos de cáncer la literatura encontrada no se basaba necesariamente en PET y se utilizaban otras técnicas que no están relacionadas con las técnicas que presentamos aquí salvo en los casos de técnicas de carácter general. Por tanto, podemos concluir que la literatura en la que nos estamos basando es de carácter específico y nosotros lo que estamos haciendo es tratar las imágenes de cuerpo completo (salvo las extremidades inferiores) de los pacientes. Entonces, es natural que sea complicado encontrar técnicas que den buenos resultados en imágenes tan complejas como las que se presentan.

Aunque no se hayan mostrado los resultados obtenidos en las técnicas de multi-umbralización que hemos probado más recientemente, son una muy buena opción para la continuación de este trabajo. En concreto, se ha probado una técnica de multi-umbralización que consiste en la simplificación del histograma mediante el uso de *Wavelet Atrous* y un posterior análisis de los máximos locales. Los resultados obtenidos no han sido buenos de cara a la delineación, pero sí se han obtenido muy buenos resultados en todos los pacientes en la separación del fondo de la imagen respecto del cuerpo del paciente. La combinación de los resultados obtenidos en esta técnica junto con la medida de *source-to-background* (S/B) ya introducida en la sección 3.4.1 permitirán acentuar el contraste de los niveles de SUV entre el cuerpo del paciente y los tumores facilitando la umbralización puesto que se estará consiguiendo lo que se buscaba desde un principio, que es la obtención de histogramas claramente bimodales. De hecho en el artículo de Erdi et al. (1997) [4] se dice explícitamente que el valor de S/B tiene en cuenta los valores de background en la definición estándar del SUV, lo cual no se ha tenido en cuenta en ninguna de estas pruebas.

Además de esta simplificación de los datos se puede llevar a cabo haciendo uso de heurísticas en base a conocimientos previos relacionados por ejemplo con la localización del tumor cancerígeno en el cuerpo. De este modo, se puede conseguir eliminar volúmenes que puedan ser complicados de discriminar tras hacer la segmentación tal y

como sucede con los riñones o volúmenes que afectan notablemente a los resultados de la umbralización como la vejiga.

Por otro lado, el uso de técnicas que no estén basadas en la umbralización no ha sido reflejado en este documento aunque eso no significa que estas técnicas no sean convenientes. Se han encontrado algunas técnicas que no han sido revisadas con suficiente detalle y, por tanto, se ha considerado mejor no describirlas. Las técnicas que pueden ser útiles para trabajos futuros son la complejidad de estos métodos encontrados han llevado al abandono de esta línea. De todas estas técnicas, podemos destacar la técnica de segmentación de PET de tipo local *fuzzy* del artículo [9] y, sobre todo, una técnica encontrada recientemente [2] que hace pruebas utilizando *phantoms* y pacientes.

Para agilizar la fase de pruebas y de discriminación de volúmenes, el desarrollo de una herramienta a parte es muy conveniente. Lo ideal es que esta herramienta sirva para facilitar la visualización y el etiquetado de estos volúmenes así como la muestra de características como por ejemplo volumen cuantitativo, el SUV mínimo y el SUV máximo. Así se podrían discriminar los volúmenes manualmente reduciendo la cantidad de tiempo requerida para dicha tarea. Alternativamente, en lugar de tratarse una herramienta manual, se puede hacer una herramienta automática que sea capaz de hacer dicha discriminación usando las características de los tumores. La elección volúmenes a discriminar no es una tarea trivial ya que, para ser exactos, el volumen más pequeño encontrado en las ROIs reales (las delineadas por los operadores o especialistas) era de 4 vóxeles. Pero eso no quiere decir que no se pueda conseguir establecer alguna relación que ayude al desarrollo de la herramienta.


Bibliografía

- [1] Isaac Bankman and Joseph Bronzino. *Handbook of Medical Imaging Processing and Analysis*, chapter II Segmentation. Academic Press, 2000.
- [2] Saoussen Belhassen and Habib Zaidi. A novel fuzzy c-means algorithm for unsupervised heterogeneous tumor quantification in pet. *Medical physics*, 37(3):1309–1324, 2010.
- [3] Kenneth J Biehl, Feng-Ming Kong, Farrokh Dehdashti, Jian-Yue Jin, Sasa Mutic, Issam El Naqa, Barry A Siegel, and Jeffrey D Bradley. 18f-fdg pet definition of gross tumor volume for radiotherapy of non-small cell lung cancer: is a single standardized uptake value threshold approach appropriate? *Journal of Nuclear Medicine*, 47(11):1808–1812, 2006.
- [4] Yusuf E Erdi, O Mawlawi, Steven M Larson, M Imbriaco, H Yeung, R Finn, and John L Humm. Segmentation of lung lesion volume by adaptive positron emission tomography image thresholding. *Cancer*, 80(S12):2505–2509, 1997.
- [5] Yusuf E Erdi, Kenneth Rosenzweig, Alev K Erdi, Homer A Macapinlac, Yu-Chi Hu, Louise E Braban, John L Humm, Olivia D Squire, Chen-Shou Chui, Steven M Larson, et al. Radiotherapy treatment planning for patients with non-small cell lung cancer using positron emission tomography (pet). *Radiotherapy and Oncology*, 62(1):51–60, 2002.
- [6] Leonard Fass. Imaging and cancer: a review. *Molecular oncology*, 2(2):115–152, 2008.
- [7] Inga S Grills, Di Yan, Quinten C Black, Ching-Yee O Wong, Alvaro A Martinez, and Larry L Kestin. Clinical implications of defining the gross tumor volume with combination of ct and^{18f}fdg-positron emission tomography in non-small-cell lung cancer. *International Journal of Radiation Oncology* Biology* Physics*, 67(3):709–719, 2007.

- [8] Yuhua Gu, Virendra Kumar, Lawrence O Hall, Dmitry B Goldgof, Ching-Yen Li, René Korn, Claus Bendtsen, Emmanuel Rios Velazquez, Andre Dekker, Hugo Aerts, et al. Automated delineation of lung tumors from ct images using a single click ensemble segmentation approach. *Pattern recognition*, 46(3):692–702, 2013.
- [9] Mathieu Hatt, Catherine Cheze le Rest, Alexandre Turzo, Christian Roux, and Dimitris Visvikis. A fuzzy locally adaptive bayesian segmentation approach for volume determination in pet. *Medical Imaging, IEEE Transactions on*, 28(6):881–893, 2009.
- [10] Xiaoyi Jiang and Daniel Mojon. Adaptive local thresholding by verification-based multithreshold probing with application to vessel detection in retinal images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 25(1):131–137, 2003.
- [11] John W Keyes. Suv: standard uptake or silly useless value? *Journal of Nuclear Medicine*, 36(10):1836–1839, 1995.
- [12] AD Nelson, M Werner-Wasik, W Choi, Y Arai, PF Faulhaber, N Ohri, JW Piper, KD Brockway, and AS Nelson. Pet tumor segmentation: multi-observer validation of a gradient-based method using a nslc pet phantom. *International Journal of Radiation Oncology* Biology* Physics*, 75(3):S627, 2009.
- [13] Ursula Nestle, Stephanie Kremp, Andrea Schaefer-Schuler, Christiane Sebastian-Welsch, Dirk Hellwig, Christian Rübe, and Carl-Martin Kirsch. Comparison of different methods for delineation of 18f-fdg pet-positive tissue for target volume definition in radiotherapy of patients with non-small cell lung cancer. *Journal of Nuclear Medicine*, 46(8):1342–1348, 2005.
- [14] Robert C Orth, Michael J Wallace, and Michael D Kuo. C-arm cone-beam ct: general principles and technical considerations for use in interventional radiology. *Journal of Vascular and Interventional Radiology*, 19(6):814–820, 2008.
- [15] Nobuyuki Otsu. A threshold selection method from gray-level histograms. *Automatica*, 11(285-296):23–27, 1975.
- [16] Arnold C Paulino, Mary Koshy, Rebecca Howell, David Schuster, and Lawrence W Davis. Comparison of ct-and fdg-pet-defined gross tumor volume in intensity-modulated radiotherapy for head-and-neck cancer. *International Journal of Radiation Oncology* Biology* Physics*, 61(5):1385–1392, 2005.
- [17] Elena Prieto, Pablo Lecumberri, Miguel Pagola, Marisol Gomez, Izaskun Bilbao, Margarita Ecay, Ivan Penuelas, and Josep M Marti-Climent. Twelve automated thresholding methods for segmentation of pet images: a phantom study. *Physics in medicine and biology*, 57(12):3963, 2012.

- [18] TW Ridler and S Calvard. Picture thresholding using an iterative selection method. *IEEE transactions on Systems, Man and Cybernetics*, 8(8):630–632, 1978.
- [19] Prasanna K Sahoo, SAKC Soltani, and AKC Wong. A survey of thresholding techniques. *Computer vision, graphics, and image processing*, 41(2):233–260, 1988.
- [20] Mehmet Sezgin et al. Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic imaging*, 13(1):146–168, 2004.
- [21] Mehmet Sezgin and Bülent Sankur. Image multithresholding based on sample moment function. In *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*, volume 2, pages II–415. IEEE, 2003.
- [22] HJ Trussell. Comments on”picture thresholding using an iterative selection method”. *Systems, Man and Cybernetics, IEEE Transactions on*, 9(5):311–311, 1979.
- [23] Koon-Pong Wong, Dagan Feng, Steven R Meikle, and Michael J Fulham. Segmentation of dynamic pet images using cluster analysis. *Nuclear Science, IEEE Transactions on*, 49(1):200–207, 2002.
- [24] Habib Zaidi and Issam El Naqa. Pet-guided delineation of radiation therapy treatment volumes: a survey of image segmentation techniques. *European journal of nuclear medicine and molecular imaging*, 37(11):2165–2187, 2010.

Este documento esta firmado por

	Firmante	CN=tfgm.fi.upm.es, OU=CCFI, O=Facultad de Informatica - UPM, C=ES
	Fecha/Hora	Fri Jun 06 22:36:41 CEST 2014
	Emisor del Certificado	EMAILADDRESS=camanager@fi.upm.es, CN=CA Facultad de Informatica, O=Facultad de Informatica - UPM, C=ES
	Numero de Serie	630
	Metodo	urn:adobe.com:Adobe.PPKLite:adbe.pkcs7.shal (Adobe Signature)